

FIGURE 1

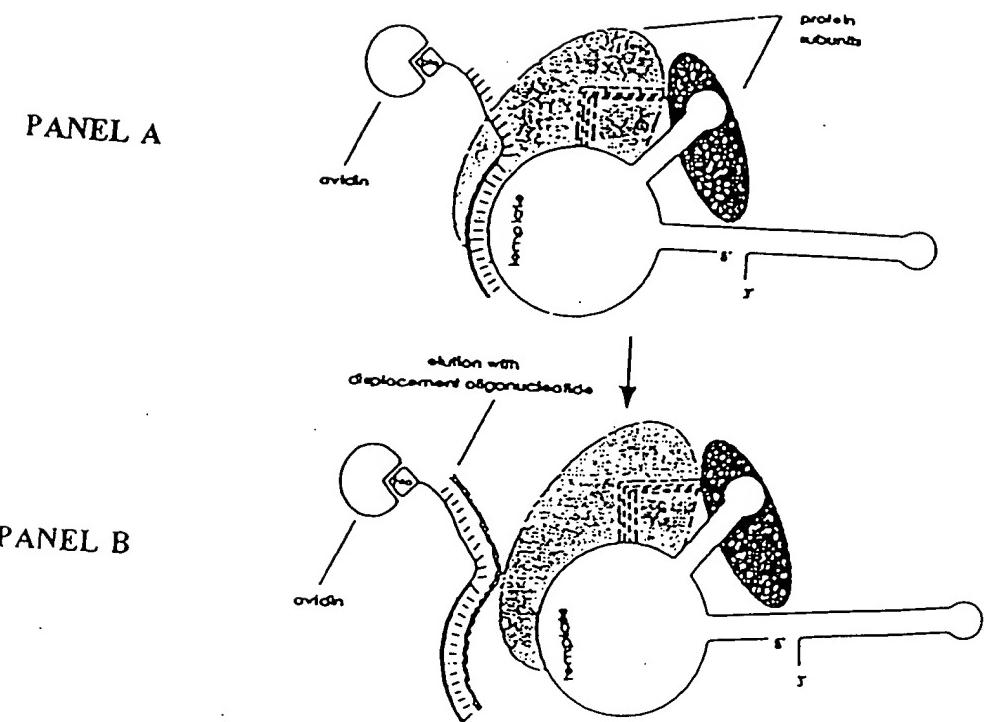


FIGURE 2

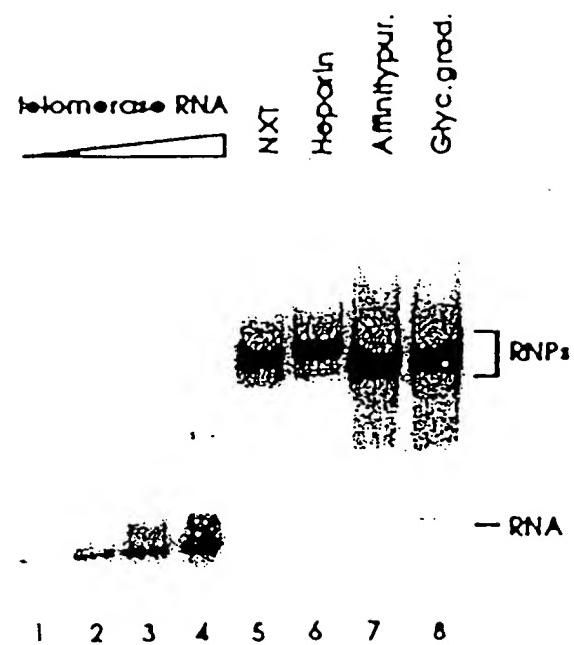


FIGURE 3

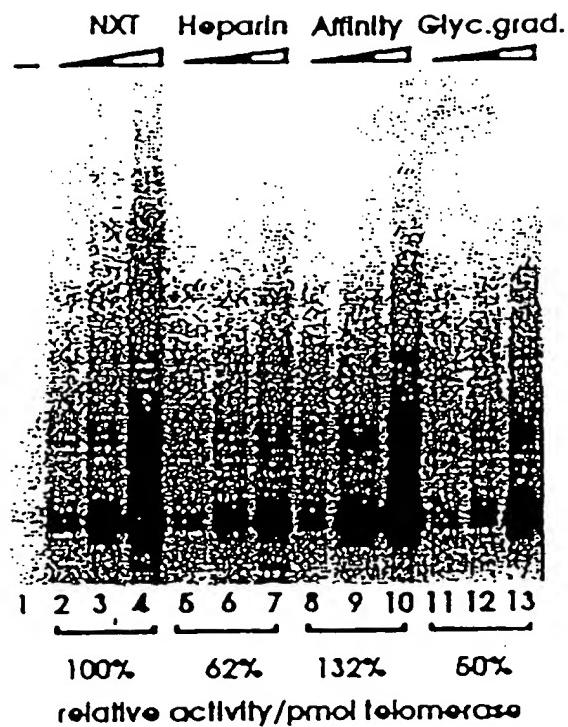


FIGURE 4

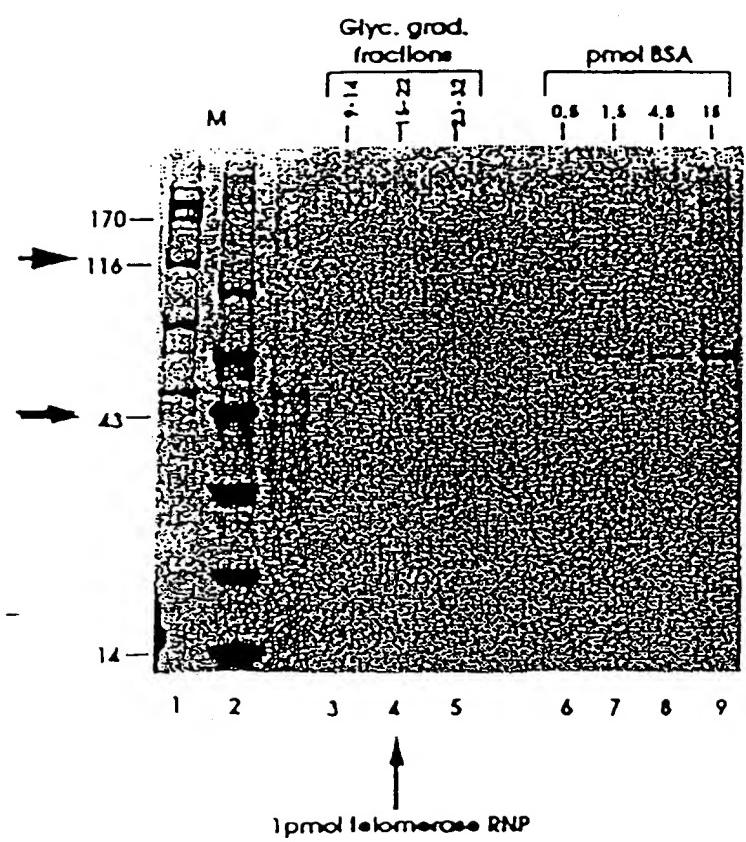


FIGURE 5

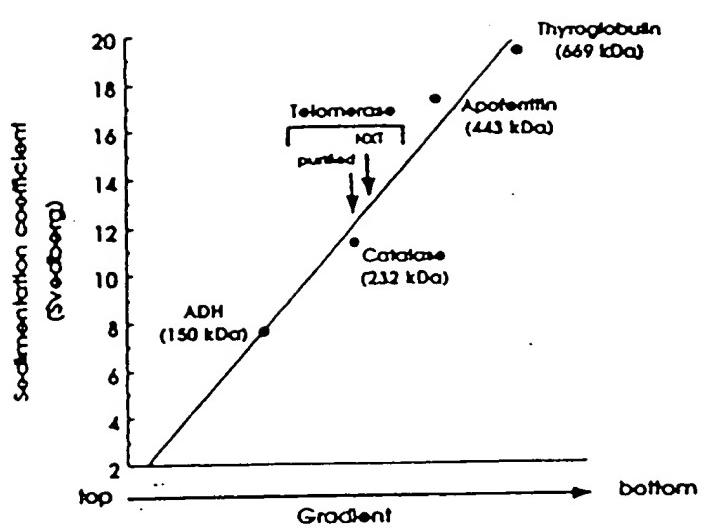


FIGURE 6

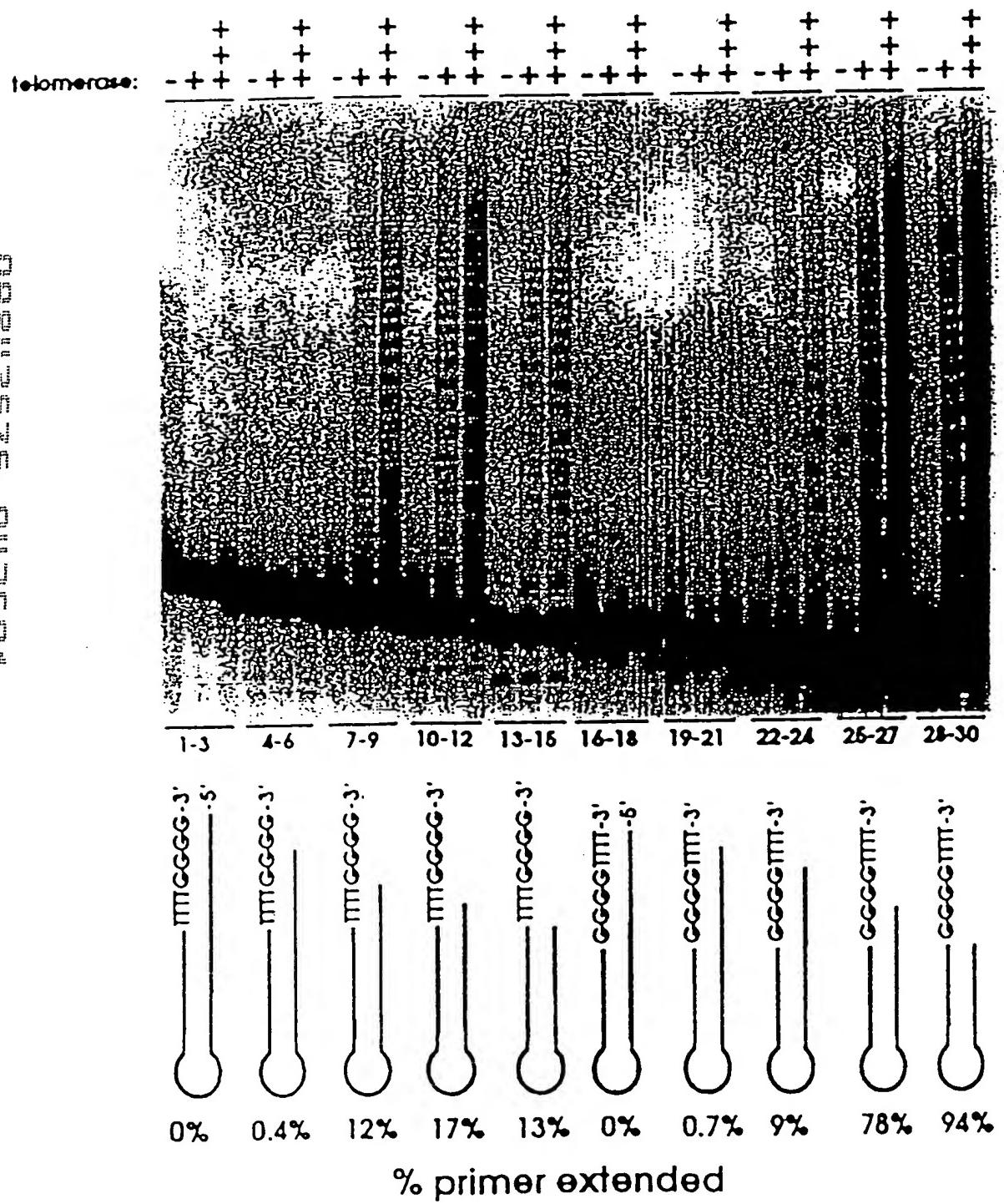


FIGURE 7

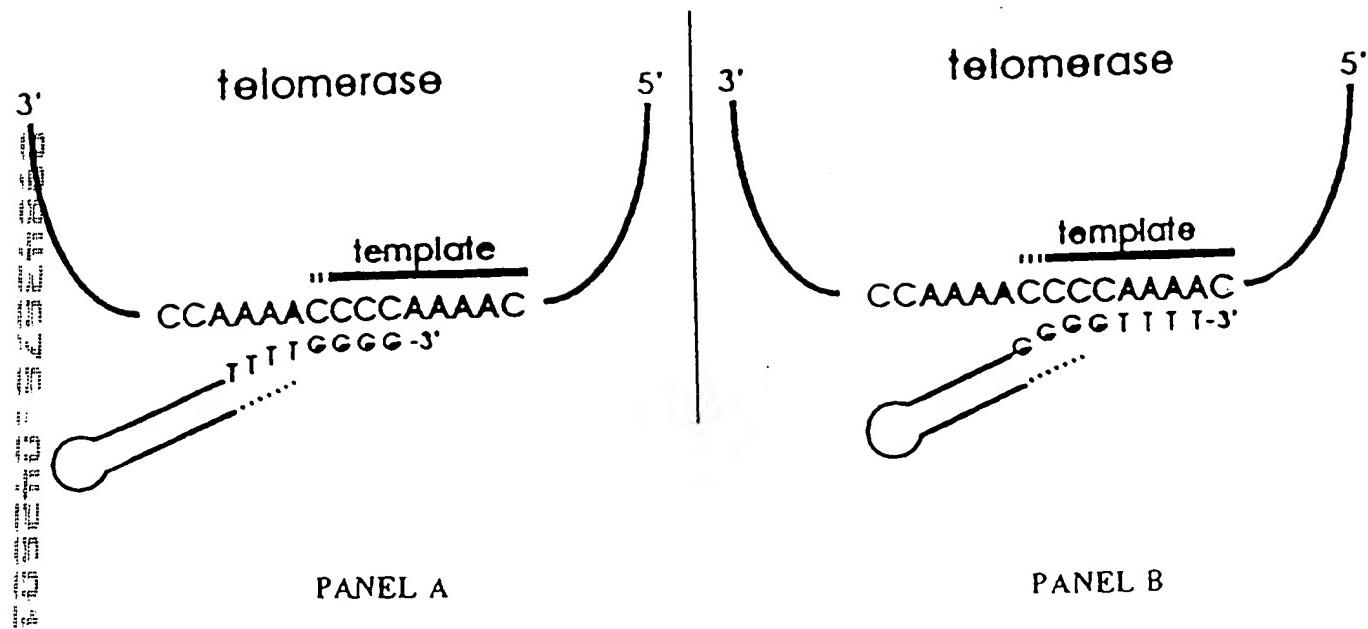


FIGURE 8

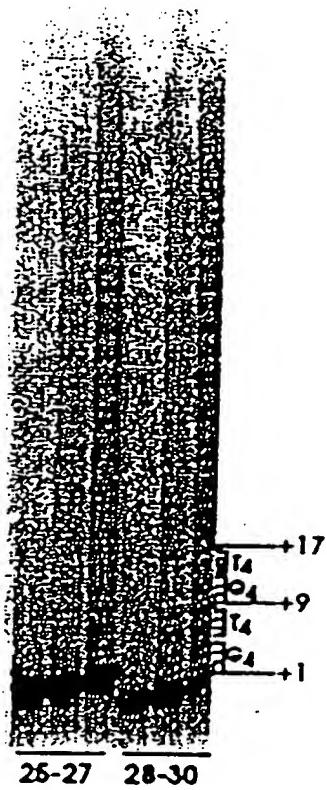


FIGURE 9

1 AAAACCCAA AACCCAAAA CCCCTTTAG AGCCCTGCAG TTGGAAATAT
51 AACCTCAGTA TTAATAAGCT CAGATTTAA ATATTAATTAA CAAAACCTAA
101 ATGGAGGTTG ATGTTGATAA TCAAGCTGAT AATCATGGCA TTCACTCAGC
151 TCTTAAGACT TGTGAAGAAA TTAAAGAAC TAAAACGTTG TACTCTTGGAA
201 TCCAGAAAGT TATTAGATGA AGAAATCAAT CTCAAAGTCA TTATAAAGAT
251 TTAGAAGATA TTAAAATATT TGCGCAGACA AATATTGTTG CTACTCCACG
301 AGACTATAAT GAAGAAGATT TTAAAGTTAT TGCAAGAAAAA GAAGTATT
351 CAACTGGACT AATGATCGAA CTTATTGACA AATGCTTAGT TGAACCTTC
401 TCATCAAGCG ATGTTTCAGA TAGACAAAAA CTTCAATGAT TTGGATTCA
451 ACTTAAGGGA AATCAATTAG CAAAGACCCA TTTATTAACA GCTCTTCAA
501 CTCAAAAGCA GTATTTCTT CAAGACGAAT GGAACCAAGT TAGAGCAATG
551 ATTGGAAATG AGCTCTTCCG ACATCTCTAC ACTAAATATT TAATATTCCA
601 GCGAACTTCT GAAGGAACCTC TTGTTCAATT TTGCAGGAAT AACGTTTG
651 ATCATTTGAA AGTCAACGAT AAGTTGACA AAAAGCAAAA AGGTGGAGCA
701 GCAGACATGA ATGAACCTCG ATGTTGATCA ACCTGCAAAT ACAATGTCAA
751 GAATGAGAAA GATCACTTTC TCAACAAACAT CAACGTGCCG AATTGGAATA
801 ATATGAAATC AAGAACCGAG ATATTTATT GCACTCATT TAATAGAAAT
851 AACCAATTCT TCAAAAAGCA TGAGTTGTG AGTAACAAAA ACAATATTTC
901 AGCGATGGAC AGAGCTCAGA CGATATTCAC GAATATATTAG ATTAAATA
951 GAATTAGAAA GAAGCTAAAAA GATAAGGTTA TCGAAAAAAAT TGCTTACATG
1001 CTTGAGAAAG TCAAAGATT TAACTTAAC TACTATTAA CAAAATCTT
1051 TCCTCTTCCA GAAAATTGGC GGGAACGGAA ACAAAAAATC GAAAACATTGA
1101 TAAATAAAAC TAGAGAAGAA AAGTCGAAGT ACTATGAAGA GCTGTTAGC
1151 TACACAACCTG ATAATAAAATG CGTCACACAA TTTATTAATG AATTTTCTA
1201 CAATATACTC CCCAAAGACT TTTTGAUTGG AAGAAACCGT AAGAATTTC
1251 AAAAGAAAGT TAAGAAATAT GTGGAACCTAA ACAAGCATGA ACTCATTAC
1301 AAAAACTTAT TGCTTGAGAA GATCAATACA AGAGAAATAT CATGGATGCA
1351 GGTTGAGACC TCTGCAAAGC ATTTTTATTA TTTTGAUTCAC GAAAACATCT
1401 ACGTCTTATG GAAATTGCTC CGATGGATAT TCGAGGATCT CGTGTCTCG
1451 CTGATTAGAT GATTTTCTA TGTCACCGAG CAACAGAAAAA GTTACTCCAA
1501 AACCTATTAC TACAGAAAGA ATATTTGGGA CGTCATTATG AAAATGTCAA
1551 TCGCAGACTT AAAGAAGGAA ACGCTTGCTG AGGTCCAAGA AAAAGAGGTT
1601 GAAGAATGGA AAAAGTCGCT TGGATTGCA CCTGGAAAAC TCAGACTAAT
1651 ACCGAAGAAA ACTACTTTCC GTCCAATTAT GACTTTCAAT AAGAAGATTG
1701 TAAATTCAAGA CCGGAAGACT ACAAAATTAA CTACAAATAC GAAGTTATTG
1751 AACTCTCACT TAATGCTTAA GACATTGAAG AATAGAATGT TTAAAGATCC
1801 TTTGGATTC GCTGTTTTA ACTATGATGA TGTAATGAAA AAGTATGAGG
1851 AGTTGTTG CAAATGGAAG CAAGTTGGAC AACCAAAACT CTTCTTGC
1901 ACTATGGATA TCGAAAAGTG ATATGATAGT GTAAACAGAG AAAAACTATC
1951 AACATTCCA AAAACTACTA AATTACTTTC TTCAAGATTTC TGGATTATGA
2001 CTGCACAAAT TCTAAAGAGA AAGAATAACA TAGTTATCGA TTGAAAAAC
2051 TTAGAAGA AAGAAATGAA AGATTATTAG AGACAGAAAT TCCAGAAGAT
2101 TGCACTTGAA GGAGGACAAT ATCCAACCTT ATTCAAGTGT CTTGAAAATG
2151 AACAAATGA CTAAATGCA AAGAAAACAT TAATTGTTGA AGCAAAGCAA
2201 AGAAATTATT TTAAGAAAGA TAACTTACTT CAACCAGTCA TTAATATTG
2251 CCAATATAAT TACATTAAC TTAATGGGA GTTTATAAA CAAACAAAAG
2301 GAATTCCCTCA AGGTCTTGA GTTTCATCAA TTTTGTCTAC ATTATTATT

FIGURE 9 (cont.)

2351 GCAACATTAG AGGAAAGCTC CTTAGGATT CTTAGAGATG AATCAATGAA
2401 CCCTGAAAAT CCAAATGTTA ATCTTCTAAT GAGACTTACA GATGACTATC
2451 TTTTGATTAC AACTCAAGAG AATAATGCAG TATTGTTTAT TGAGAAACTT
2501 ATAAACGTAA GTCGTAAAAA TGGATTTAAA TTCAATATGA AGAAACTACA
2551 GACTAGTTT CCATTAAGTC CAAGCAAATT TGCAAAATAC GGAATGGATA
2601 GTGTTGAGGA GCAAAATATT GTTCAAGATT ACTGCGATTG GATTGGCATC
2651 TCAATTGATA TGAAAACCTCT TGCTTTAATG CCAAATATTA ACTTGAGAAT
2701 AGAAGGAATT CTGTGTACAC TCAATCTAAA CATGCAAACA AAGAAAGCAT
2751 CAATGTGGCT CAAGAAGAAA CTAAAGTCGT TTTTAATGAA TAACATTACC
2801 CATTATTTA GAAAGACGAT TACAACCGAA GACTTGCAG ATAAAACCTCT
2851 CAAACAAGTTA TTTATATCAG GCGGTTACAA ATACATGCAA TGAGCCAAAG
2901 AATACAAGGA CCACCTTAAG AAGAACCTAG CTATGAGCAG TATGATCGAC
2951 TTAGAGGTAT CTAAAATTAT ATACTCTGTA ACCAGAGCAT TCTTTAAATA
3001 CCTTGTGTGC AATATTAAGG ATACAATTIT TGGAGAGGGAG CATTATCCAG
3051 ACTTTTCCT TAGCACACTG AAGCACTTA TTGAAATATT CAGCACAAAA
3101 AAGTACATT TCAACAGAGT TTGCATGATC CTCAAGGCAA AAGAAGCAAA
3151 GCTAAAAAGT GACCAATGTC AATCTCTAAT TCAATATGAT GCATAGTCGA
3201 CTATTCTAAC TTATTTGGA AAGTTAATT TCAATTITG TCTTATATAC
3251 TGGGGTTTTG GGGTTTGGG GTTTGGGG

FIGURE 10

1 MEVDVDNQAD NHGIHSALKT CEEIKEAKTL YSWIQQVIRC RNQSQS HYKD
51 LEDIKIFAQT NIVATPRDYN EEDFKVIARK EVFSTGLMIE LIDKCLVELL
101 SSSDVSDRQK LQCFGQLKG NQLAKTHLLT ALSTQKQYFF QDEWNQVRAM
151 IGNELFRHLY TKYLIFQRTS EGTLVQFCGN NVFDHLKVND KFDKKQKGGA
201 ADMNEPRCCS TCKYNVKNEK DHFLNNINVP NWNNMKSRTTR IFYCTHFNRN
251 NQFFKKHEFV SNKNNISAMD RAQTIFTNIF RFNRIRKKLK DKVIEKIAYM
301 LEVKVDFNFN YYLTOKSCPLP ENWRERKQKI ENLINKTREE KSKYYEELFS
351 YTTDNKCVTQ FINEFFYNIL PKDFTGRNR KNFQKKVKKY VELNKHELIH
401 KNLLLEKINT REISWMQVET SAKHFYYFDH ENIYVLWKLL RWIFEDLVVS
451 LIRCFFYVTE QQKSYSKTYY YRKNIWDVIM KMSIADLKKE TLAEVQEKEV
501 EEWKKSLGFA PGKLRLIPKK TTFRPIMTFN KKIVNSDRKT TKLTTNTKLL
551 NSHMLKTLK NRMFKDPFGF AVFNYDDVMK KYEEFVCKWK QVGQPKLFFA
601 TMDIEKCYDS VNREKLSTFL KTTKLSSDF WIMTAQILKR KNNIVIDSKN
651 FRKKEMKDYF RQKFQKIALE GGQYPTLFSV LENEQNDLNA KKT LIVEAKQ
701 RNYFKKDNLQ QPVINICQYN YINFNGKFYK QTKGIPQGLC VSSILSSFYY
751 ATLEESSLGFLRDESMNPEN PNVNLLMRLT DDYLLITTQE NNAVLFIEKL
801 INVSRENGFK FNMKKLQTSF PLSPSKFAKY GMDSVEEQNI VQDYCDWIGI
851 SIDMKTALM PNINLRIEGI LCTLNLMQT KKASMWLKKK LKSFLMNNIT
901 HYFRKTITTE DFANKTLNKL FISGGYKYMQ CAKEYKDHFK KNLAMSSMID
951 LEVSKIYSV TRAFFKYLVC NIKDTIFGEE HYPDFFLSTL KHFIEIFSTK
1001 KYIFNRVCMI LKAKEAKLKS DQCQSLIQYD A

FIGURE 11

1 CCCCAAAACC CCAAAACCCC AAAACCCCTA TAAAAAAAAGA AAAAATTGAG
51 GTAGTTAGA AATAAAATAT TATTCCCGCA CAAATGGAGA TGGATATTGA
101 TTTGGATGAT ATAGAAAATT TACTTCCTAA TACATTCAAC AAGTATAGCA
151 GCTCTTGTAG TGACAAGAAA GGATGCAAAA CATTGAAATC TGGCTCGAAA
201 TCGCCTTCAT TGACTATTCC AAAGTTGCAA AAACAATTAG AGTTCTACTT
251 CTCGGATGCA AATCTTTATA ACGATTCTT CTTGAGAAAA TTGTTTAA
301 AAAGCGGAGA GCAAAGAGTA GAAATTGAAA CATTACTAAT GTTTAAATAA
351 AATCAGGTAA TGAGGATTAT TCTATTTTT AGATCACTTC TTAAGGAGCA
401 TTATGGAGAA ATTACTTAA TACTAAAAGG TAAACAGTTT GGATTATTC
451 CCTAGCCAAC AATGATGAGT ATATTAAATT CATATGAGAA TGAGTCAAAG
501 GATCTCGATA CATCAGACTT ACCAAAGACA AACTCGCTAT AAAACGCAAG
551 AAAAAGTTTG ATAATCGAAC AGCAGAAGAA CTTATTGCAT TTACTATTG
601 TATGGGTTTT ATTACAATTG TTTAGGTAT CGACGGTGAA CTCCCGAGTC
651 TTGAGACAAT TGAAAAAGCT GTTACAAC GAAGGAATCG CAGTTCTGAA
701 AGTTCTGATG TGTATGCCAT TATTTGTGA ATTAATCTCA AATATCTTAT
751 CTCATTAA TGAGATGCTA TAGAAACAAA CCAAATAAAC CATGCAAGTT
801 TAATGGAATA TACGTTAAAT CCTTTGGGAC AAATGCACAC TGAATTATA
851 TTGGATTCTT AAAGCATAGA TACACAGAA GCTTTAGAGA CTGATTAGC
901 TTACAACAGA TTACCTGTTT TGATTACTCT TGCTCATCTC TTATATCTT
951 AAAAGAAGCA GGCGAAATGA AAAGAAGACT AAAGAAAGAG ATTICAAAAT
1001 TTGTTGATTTC TTCTGTAACC GGAATTAACA ACAAGAATAT TAGCAACGAA
1051 AAAGAAGAAG AGCTATCACA ATCCTGATTTC TTAAAGATTT CAAAAATTCC
1101 AGGTAAAGAGA GATACATTCA TTAAAATTCA TATATTATAG TTTTCATT
1151 CACAGCTGTT ATTTCTTTT ATCTTAACAA TATTTTTGAA TTAGCTGGAA
1201 GTAAAAAGTA TCAAATAAGA GAAGCGCTAG ACTGAGGTAA CTTAGCTTAT
1251 TCACATTCA AGATCGACCT TCATATATCC AATACGATGA TAAGGAAACA
1301 GCAGTCATCC GTTTAAAAAA TAGTGTATG AGGACTAAAT TTTTAGAGTC
1351 AAGAAATGGA GCCGAAATCT TAATCAAAAAA GAATTGCGTC GATATTGCAA
1401 AAGAATCGAA CTCTAAATCT TTCGTTAATA AGTATTACCA ATCTTGATTG
1451 ATTGAAGAGA TTGACGAGGC AACTGCACAG AAGATCATTAA AAGAAATAAA
1501 GTAACTTTA TTAATTAGAG AATAAACTAA ATTACTAATA TAGAGATCAG
1551 CGATCTCAA TTGACGAAAT AAAAGCTGAA CTAAAGTTAG ACAATAAAAA
1601 ATACAAACCT TGGTCAAAAT ATTGAGGAAG GAAAAGAAGA CCAGTTAGCA
1651 AAAGAAAAAA TAAGGCAATA AATAAAATGA GTACAGAAGT GAAGAAATAA
1701 AAGATTATT TTTTCAATA ATTTATTGAA AAGAGGGGTT TTGGGGTTTT
1751 GGGGTTTGG GG

FIGURE 12

1 CCCCCAAAACCCAAAAACCCAAAACCCCTATAAAAAAGAAAAAAATTGAGGTAGTTAGA 60
 1 GGGGTTTTCGGGTTTGGGGTTTGGGGATATTTTTCTTTTAACTCATCAAATCT
 a P Q N P K T P K P L * K K K K L R * F R -
 b P K T P K P Q N P Y K K R K N C G S L E -
 c P K P Q N P K T P I K K E K I E V V * K -
 AATAAAATATTATTCCCGACAATGGAGATGGATATTGATTTGGATGATAGAAAATT 120
 TTATTTATAATAAGGGCGTGTACCTCACCTATAACTAAACCTACTATATCTTTAA
 a N K I L F P H K W R W I L I W M I * K I -
 b I K Y Y S R T N G D G Y C F G C Y R K F -
 c * N I I P A Q M E M D I D L D D I E N L -
 TACTTCCTAACATTCAACAAGTATAGCAGCTTGAGACAGAAAGGATGCAAAA 180
 ATGAAGGATTATGTAAGTTGTTCATATCGTCGAGAACATCACTGTTCTTCTACGTTT
 a Y F L I H S T S I A A L V V T R K D A K -
 b T S * Y I Q Q V * Q L L * Q E R M Q N -
 c L P N T F N K Y S S S C S D K K G C K T -
 CATGAAATCTGGCTCGAAATCGCTTCATTGACTATTCAAAGTTGCAAAAACATTAG 240
 GTAACTTTAGACCGAGCTTAGCGGAAGTAACTGATAAGGTTCAACGTTTGTAACT
 a H C N L A R N R L H C L F Q S C K N N -
 b I E I W L E I A F I D Y S K V A K T I R -
 c L K S G S K S P S L T I P K L Q K Q L E -
 AGTTCTACTTCGGATGCAAATC'TTTATAACGATTCTTCTTGAGAAAATTAGTTAA 300
 TCAAGATGAAGAGCCTACGTTAGAAATATTGCTAAGAAAGAACTTTAAATCAAATT
 a S S T S R M Q I F I T I L S C E N * F -
 b V L L E G C K S L * R F F L E K I S F K -
 c F Y P S D A N L Y N D S F L R K L V L K -
 AAAGCCGAGAGCAAAGAGTAGAAATTGAAACATTACTAATGTTAAATAAAATCACCTAA 360
 TTTGGCTCTGGTTCTCATTTTAACCTTCTAATGATTACAATTATTTACTCCATT
 a K A E S K E * K L K H Y * C L N K I R -
 b K R R A K S R N C N I T N V * I K S G N -
 c S G E Q R V E I E T L L M P K * N Q V M -
 TGACGGATTATTCTATTTCAGATCACTTCAGGAGCATTATGGAGAAAATTACTTAA 420
 ACTCCTAACATAAGATAAAAATCTAGTAAGAATTCCCTGTAATACCTTTAAATGAAATT
 a C G L F Y F L D H F L R S I M E K I T -
 b E D Y S I F * I T S * G A L W R K L L N -
 c R I I L F F R S L L K E H Y G E N Y L I -
 TACTAAAGGTAAACAGTTGATTATTCCTAGCCAACAATGATGAGTATATTAAATT 480
 ATGATTTCCATTGTCAAACCTAACAAAGGGATCGCTTCTACTACTCATATAATTAA
 a Y * K V N S L D Y F P S Q Q C C V Y * I -
 b T K R * T V W I I S L A N N D E Y I K F -
 c L K G X Q F C L F P * P T M M S I L N S -

FIGURE 12 (cont.)

481 CATATGAGAATGACTCAAAGGATCTCGATAACATCAGACTTACCAAGACAAACTCGCTAT 540
 GTATACTCTTAACCTAGTTCCTAGAGCTATCTAGTCTGAATGGTTCTGTTGACCGATA
 a H M R M S Q R I S I H O T Y Q R Q T R Y -
 b I C E C V K G S R Y I R L T K D K L A I -
 c Y E N E S K D L D T S D L P K T N S L -
 541 AAAACGCAAGAAAAAGTTGATAATCGAACAGCAGAAGAACTTATTGCATTTACTATTGC 600
 TTTTCCGTTCTTTCAAACTATTAGCTTGTCTTCTGAATAACGTAAATGATAAGC
 a K T Q E K V C * S N S R R T Y C I Y Y S -
 b K R K K K F D N R T A E E L I A F T I R -
 c N A R K S L I I E Q Q K N L L H L L F V -
 601 TATGGGTTTATTACAATTGTTTAGGTATCGACGGTGAACCTCCGAGTCTTGAGACAAT 660
 ATACCCAAAATAATGTTAACAAATCCATAGCTGCCACTTGAGGGCTCAGAACTCTGTTA
 a Y G F Y Y N C F R Y R R C T P E S C D N -
 b M G F I T I V L G I D G E L P S L E T I -
 c W V L L Q L F * V S T V N S R V L R Q L -
 661 TGAAGGACTGTTACAACCTAGGAATCGCAGTTCTGAAAGTTCTGATGTGTATGCCAT 720
 ACTTTTCGACAAATGTTGACTTCCCTAGCGTCAAGACTTCAAGACTACACATACCGTA
 a C K S C L Q L K E S Q F C K F * C V C H -
 b E K A V Y N C R N R S S E S S D V Y A I -
 c K K L F T T E G I A V L K V L H M C M P L -
 721 TATTTTGTAATTAACTCAAATATCTTATCTCAATTAAATGGATAGCTATAGAAACAAA 780
 ATAAAACACTTAATTAGAGTTATAGAATAGAGTTAAATTACCTATCGATATCTTGT
 a Y F V N * S Q I S Y L N L M D S Y R N K -
 b I L C I N L K Y L I S I * W I A I E T N -
 c F C E L I S N I L S Q F N G * L * K Q T -
 781 CCAAAATAACCATCCAAGTTAACGAAATACGTTAACATCTTGGGACAAATGCACAC 840
 CCTTTATTTGGTACGTTCAATTACCTTATATGCAATTAGAAACCTGTTACGTGTG
 a P N K P C K F N G I Y V K S F G T N A H -
 b O I N H A S L M E Y T L N P L C Q M H T -
 c K * T X Q V * W N I R * I L W O K C T L -
 841 TGAATTATATGGATTCTTAAAGCATAGATAACACAGAAATGCTTTAGAGACTGATTAGC 900
 ACTTAAATATAACCTAACGAAATTCTGATCTATGTGTCTTACGAAATCTGACTAAATCG
 a C I Y E C F L K H R Y T E C F R D C F S -
 b E F I D S * S I D T Q N A L E T D L A -
 c N L Y W I L K A * I H R M L * R L I * L -
 901 TTACAAACAGATTACCTGTTGATTACTCTGCTCATCTCTTATATCTTTAAAAGAGCA 960
 AATGTTGTCTAACGACAAAACAAATGAGAACGAGTAGAGAAATAGAAATTCTTCTCGT
 a L Q Q I T C F D Y S C S S L I S L K E A -
 b Y N R L P V L I T L A H L L Y L * K K Q -
 c T T D Y L F C L L L L I S Y I F K R S R -
 961 GCGCAAATGAAAAGAACACTAAAGAAAGAGATTCTAAACCTTGTGATTCTCTGTAACC 1020
 CCCCTTTACTTTCTGATTCTTCTCTAAAGTTAACAAACTAACGAAAGACATTGG
 a G E M K R R L K K E I S K F V D S S V T -
 b A K C K E D * R K R F Q N L L I L L * P -
 c R N E K X T S E R D F K I C C F F C N R -
 1021 GGAATTAACAAAGAAATATTAGCAACGAAAAAGAAGAACGACTACAAATCCTGATTG 1080
 CCTTAATTGTTCTTATAATCCTGCTTTCTCTGATAGTGTAGGACTAAAG
 a G I N N K N I S N E K E E E L S O S C F -
 b E L T T R I L A T K K K K S Y H N P D S -
 c N * Q Q E Y * Q R K R R A I T I L I L -

FIGURE 12 (cont.)

TAAAGATTTCAAAATTCCAGGTAAGAGAGATACTATCATTAAATTATATAG
 1081 AATTCTAAAGTTAAGGCCATTCTCTATGTAAGTAAATTAAATATC
 L K I S K I P G K R D T F I K I H I L -
 . R F Q K F Q V R E I H S L K F I Y Y S -
 K D F K N S R . E R Y I H . N S Y I I V -
 TTTTCATTTACAGCTGTTATTTCTTTATCTAACAAATATTTTGATTAGCTGGAA
 1141 AAAAGTAAAGTGTGACAATAAAAAGAAAATAGAATTGTTATAAAAAACTAATCGACCTT
 F F I S Q L L F S F I L T I F F D . L E -
 F S F H S C Y F L L S . Q Y F L I S W K -
 F H F T A V I F F Y L N N I F C L A G S -
 GTAAAAAGTATCAAATAAGAGAAGCGCTAGACTGAGCTAACCTAGCTTATTCACATTCA
 1201 CATTTTCATAGTTATTCCTCGCATCTGACTCCATTGAATCGAATAAGTGTAAAGTA
 V K S I K . E K R . T E V T . L I H I H -
 . K V S N K R S A R L R . L S L F T F I -
 K K Y Q I R E A L D C G N L A Y S H S -
 AGATCGCACCTTCATATATCCAATACGATGATAACGAAACAGCAGTCATCCGTTTTAAAAA
 1261 TCTAGCTGGAAGTATATAGCTTATGCTACTATTCCCTTGTGTCAGTAGGCAAAATTTT
 R S T F I Y P I R C . G N S S H P F . K -
 D R P S Y I Q Y D D K E T A V I R F K N -
 I D L H I S N T M I R K Q O S S V L K I -
 TAGTGCTATGAGGACTAAATTTAGAGTCAGAAATGGAGGCCAAATCTTAATCAAAAA
 1321 ATCACGATACTCCTGATTTAAAATCTCAGTTCTTACCTCGGCTTCTAGAATTAGTTTT
 . C Y E D . I F R V K K W S R N L N Q K -
 S A M R T K F L E S R N G A E I L I K K -
 V L C G L N F . S Q E M E P K S . S K R -
 GAATTCCGTCGATATTCAAAGAATCGAACTCTAAATCTTCGTTAAATAAGTATTACCA
 1381 CTTAACCGCAGCTATAACGTTCTTAGCTTGTAGATTAGAAAGCAATTATTCTATAATGGT
 E L R R Y C K R I E L . I F R . . V L P -
 N C V D I A K E S N S K S F V N K Y Y Q -
 I A S : L Q K N R T L N L S L I S I T N -
 ATCTTGATTGATTGAAGAGATTGACCGAGCCAACTGCACAGAAGATCTAAAGAAATAAA
 1441 TAGAACTAACTTCTCTACTGCTCCGTTGACCTGCTCTAGTAATTCTTATTT
 I L I D C R D . R G N C T E D H . R N K -
 S C L E E E I D E A T A Q K I I K E I K -
 L D C L K R L T R Q L H R R S L K K . S -
 CTAACCTTTATTAAATTAGAGAATAAAACTAAATTACTAATATAGAGATCAGCGATCTCAA
 1501 CATTGAAATTAATCTCTTATTTGATTAAATGATTATCTCTAGTCGCTAGAAGTT
 V T F I N . R I N . I T N E E I S D L Q -
 . L L L I R E . T K L L I . R S A I F N -
 N F Y . L E N K L N Y . Y R D Q R S S I -
 TTGACGAAATAAAAGCTGAACCTAAAGTTAGACAAATAAAACCTGGCTCAAAAT
 1561 AACTGCTTATTTGACTTCAATCTGTTATTTTATCTTGGACCAACAGTTTA
 L T K . K L N . S . T I K N T N L G Q N -
 C R N K S C T K V R Q . K I O T L V K I -
 D E I K A E L K L D N K K Y K P W S K Y -
 ATTGAGGAAGGAAAAGAAGACCACTTAGCAAAAGAAAAATAAGCCATAAAATAATG
 1621 TAACTCCTCCCTTCTCTGGTCAATCGTTCTTATTTTACTCCTTATTTTACT
 I E E G K E D O L A K E K I R O . ! K C -
 L R K E K K T S . Q K K K . G N K . N E -
 C G R K R R P V S K R K N K A I N K M S -

FIGURE 12 (cont.)

1681 GTACAGAAGTGAAGAAATAAAAGATTATTTTTCAATAATTTATTGAAAAGAGGGGTT
-----+-----+-----+-----+-----+-----+-----+
CATGTCTTCACTTCTTATTTCTAAATAAAAAAACTTATTAAATAACTTTCTCCCCAA

a V Q K C R N K R F I F F N N L L K R G V -
b Y R S E E I K D L F F S I I Y C K E G F -
c T E V K K * K I Y F F Q * F I E K R G F -

1741 TTGGGGTTTGGGGTTTGGGG
-----+-----+-----+-----+-----+-----+-----+
AACCCCAAAACCCCAAAACCC

a L G F W G F G -
b W G F G V L G -
c G V L G F W -

FIGURE 13

2 EVDVQHOADNHGIHSALKTCEEIKEAKTLYSWIQLVKVIRCRNQSOSHVKDL 51
 19 ELELEMENQNDIQVRVK...IDDPKQY..LNVNTAACLLQEGSYYQDK 62
 52 EDIKIFAOVNIVATPRDYNEEDFKVIARKEVF..STGLMIELIDKCLVELL 100
 63 DERRITTKALL...EVAESDPEFICQLAVYIRNELYIRTTNYIVAF. 107
 101 SSSDVSDRQKLOCFGFOLKGNLAKTHLLTALSTQKQYFFQDEWNQVRAM 150
 108 CVVHKNTOPFIEKYFNKAVLLPNDLLEVCEFAQVLYI 144
 151 IGNELFRHLYTKYLIFORTSEGTIVQFCGNNFQDHLKVNDFDKKKQKGG 200
 145 FDATEFKNLY...LDRILSQDIRKELTFRKCLQRCVRSKF 181
 201 ADMNE..PRCCSTCKYNVKNEKDHFLLNNINVPNWNNMKSRTIFYCTHF 247
 182 SEFNEYOLGKYCTES..QRKKTMFRLSVTNKQWDQTKKK..... 220
 248 NRNNOFFKKHEFVSNKNISAMDRAQTIFTNIFRFNRIRKKLKDVKIEKI 297
 221 RKENLLTKLQAIKESEDSKRETG...DIMNVEDAIKALKPAVMKKI 264
 298 AYMLEKVVKDFNFNYYLTKSCPLPENWRERKQKIEALKTREESKYYEE 347
 265 AKRQNAMK.....KHMKAIPNSTLESKYLTTFKD 294
 348 LFSYTTDNKCVTQFINEFFYNILPKDFLTGRNRKNFQKKVKKYVELNKHE 397
 295 LIKFCHISEP....KERVYKILGKKYPKTEEEYKAAGFDSASAPFN.PE 338
 398 LIHKONLLEKINTREISWMQVETSAKHFYYFDHENIYVLWKLLRWIFEDL 447
 339 LAGKRMKIEISKWTWENELSAKGNTAEVWDNLISSNQLPYMAMLRNLSN.. 386
 448 VVSLIRCFFYVTQQKSYSKTYYRKNIWDVIMKMSIADLKETLAEVQE 497
 387ILKAGVSD..... 394
 498 KEVEEWKKSLGFAPGKLRLIPKKTTFRPIMTFNKKIVNSDRKTTKLTNT 547
 395TTHS 398
 548 KLLNSHLMUKTLKNRMFKDPFGFAVFNYYDDVMKKYEEFVCKWKQVGQPKL 597
 399 IVINK.....ICEPKAVENSKM 415
 598 FFATMDIEKYDSVNREKLSTFLKTTKLLSSDFWIMTAQILKRKNNIVID 647
 416 F..PLOFFSATEAVN.EAVTKGFKAKK...RENMMNLKGQIEAVKE..VVE 457
 648 SKNFRKKEMKDYFRQKFQKIALEGQGYPTLFSVLENEQNDLNAAKTLIVE 697
 458 KTDEEKKD.....ELEQTEEGEFVKVNEGIGKQYINSIELAIK 496
 698 AKQRNYFKKDNLLOPVINICQNYINFNGKFYKQTGKIPQGLCVSSILSS 747
 497 IAVNKNLDEIKGHTAIFSDVSGSMSTSMSGGAKKYGSVRTCLECALVGL 546
 748 FYYATLEESSLGFLRDESMNPENPNVNLMMRLTDYLLITTOENNVALFI 797
 547 MVQRCEKSSFYIFSSPSSQCNKCYLEVDL..... 576
 798 EKLINVSRENGFKFMK..LQTSFPLSPSKFAKYGMDSVEEQNIVQDYCD 846
 577 ..PGDELRPSMOKLLQEKGKLG..TDFPYECIDEWTQNKTHVD 617
 847 WIGISIDMKTALMPNINLRIEGILCTLNLNMQTKKASMWLKKLKSFLM 896
 618 NIVILSDMMAEGYSDINVRGSSIVNSI.....KKYKDEVN 653
 897 NNITHYFRKTITTEDFANKTLNKLFISSGGYKYMQCAKEYKD.HFKKNLAM 945
 654 PNIKIF..AVDLEGYG.....KCLNLGDEFNENNYIKIFGM 687
 946 SSMIDLEVKIIYSVTRAFFKYLVCNIKDTIFGEEHYPDFFLSTLKHFI 995
 688 SDSI.....LKFISAKQGGA.....NMVE 706
 996 IFSTKKYIFNRVC 1008
 707 VI..KNFALQKIG 717

FIGURE 14

132 LSTQKQYFFQDEWNQVRAMIGNEL..FRHLYTKYLIFORTSE..GTLVQFC 178
 1 MSRRNQ.....KKPQAPIGNETNLDFVLQNLEVYKSQIEHYKTQQQO 43
 179 GNNVFDHLKVNDKFDDKKQKGGAADMNEPRCCSTCKNVKNEKDHFLLNNIN 228
 44 KEEDLKLKFKNQDQDGNSGNDODDEE.....NNSNKQQELLRRVN 84
 229 VPNWNNMKSRTRIFYCTHFNRRNNOFFKKHEFVSNNISAMDRAQTIFTN 278
 85QIKQQVQLIKK...VGSKVEKDLNLNEDENKKN 114
 279 IFRFNIRKKLKDKVIEKIAYMLEKVKDFNFNYYLTKSCPLPENWRERKQ 328
 115 GLSEQQVKEEQLRTITEEQVQVQYQNLVFNMDYQLDLNESGGHRRRRETDY 164
 329 KIENLINKTREEKSKYYEELFSYTTDNKCVTQFINE..FFYNILPKDFLTG 377
 165 DTEKWFEISHDQK.....NYSVIYANQKTSYCWWLKDYFNK 200
 378 RNRINFQKKVKYVELNKHELIHKNLLLEKINTREISWMQVETSAKHFYY 427
 201 NNYDHLNVSINRL..TEAEFYAFDDFSQTIKLTNNSYQTVNID..... 242
 428 FDHENIYVWLKLLRWI..FEDLVVSLIRCFFYVTEQQKSYSKTYYRKNI 475
 243 VNFDDNNICILALLRFLLSLERFNILNIRSSY..TRNQYNPEKIGELLETI 290
 476 WDVIHKMSIADLKETLAEVQEKEVEEWKSLGFAPGKLRILPKKTFRP 525
 291 FAVVFSHR.....HLQGIHLQVPCEAFQYLVNSSSQISVKDSQLQ 330
 526 IMTFNKKIVNSDRKTTKLTTNTKLLNSHMLKTLKNRMFKDPFGFAVFNY 575
 331 VYSFSTDLKVD..TNKVQDYFKFLQEFPRLTHSVQQAIPVSATNAVENL 378
 576 DDVHKKYEEFVKWKQVGQPKLF.....FATMDIEKCYDS..VNREK 615
 379 NVLLKKVKH..ANLNLVSIPTQFNFDYFVNLOHLKLEFGLEPNILTQKQ 426
 516 LSTFL..KTTKLLSSDFWIMTAQILKRKNNI..VIDSKNFRKEMK 657
 427 LENLLLSIKOSKNLKFRLRNFTYVAQETSRKOILKQATTIKNLKNNKNO 476
 558 DYFRQKFQKALEGGQYPFLFSVLEN..EQNDLNAAKTLIVEAKQRNYFK 705
 477 EETPETKDETPESTSGMKFFDHLSELTELEDFSVN...LQATQEYI.. 520
 706 KDNLQPVNICOYNYINFNGKFQTKGIPQGLCVSSILSSFYYATLEE 755
 521 DSHKLLIRSTNLKKFKLSKYEMEKSKMOTFIDLKNI.....YETLNN 564
 756 SSLGFLRDESHPENPNVNLMLRLTDDYLLITTOENNNAVLIEKLINVSR 805
 565 ..LKRCSVNISNPNGNISYELTN.....KDSTFYKFKLTLNQE 500
 806 ENGFKNMKKLQTSFPLSPSKFAKYGMDSVEEQNIVQDYCDWIGISIDMK 855
 501 LQHAKYTFK..ONEFQFNNVKSAKIESSSLESLEDIDSLCKSIASCKNQ 648
 856 TLALMPNINLRIEGILCTLNLMOT..KKASHWLKK..KLKSFLMNNITH 901
 649 NVNI....IASLLYPNIQKNPFNKPNNLFFKQFEQLKNLENVSINC 691
 902 YFRKTI..TTEDFANKTLNKLFISSGGYKYMOCAKEKEYKDHFKXNLAMSSM 948
 692 ILDOHQHILNSISEFLEKNKKIKAFILKRYLLQYYLDYTKLFKTLQQLPEL 741
 949 IDLEVSKIIYSVT.....RAFFKYLVCNIKDT..IFGEEHY 982
 742 NOVYINQQLEELTVSEVHKQWENHKQKAFYEPCLCEFIKESSOTLOLIDF 791
 983 PDFFLS..TLXHFILEIFSTKYY IFNRVCHMILKAKEAKLKSQDCQSLIQ 1028
 792 DONTVSDDSIKKILESISSESKYHHYRLNPSQSSSLIKSENEEIQELLK 840

FIGURE 15

4 DIDLDIENLLPNTFNKYSSSCSDKGCKTLKSGSKSPSLTIPK..... 47
617 NVKSAKIESSSLESLEDEDISLCKSIASCKNLQNVNIIASLLYPNNIQKNP 666
48LQKQLEFYFSANLYNDSFLRKVLVLSGEQRVE....IETLLM 86
667 FNKPNNLFFFQFEOLKNLENVSINCILDQHILNSISEFLEKNNKKIKAFIL 716

FIGURE 16

1 MEHDIDLODIEENL.....LPNTFNKYSSSCSDKGCKTLKSGSKSPS... 42
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
491 TELAIKIAVNKNLDEIKGHTAIFSDVSGSMSTSHSGGAKKYGSVRTCLEC 540
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
493 LTIPKLOKO.....LEFYFSANLYNDSFLRKLVLKSGEQRVEIETLL 85
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
541 ALVLGLMVKQRCEKSSFYIFSSPSSQCNKCYL.EVDLPGDELRPSHQKLL 589

FIGURE 17

	Motif A	Motif B	Motif C	Motif D	Motif E
Consensus	h---h---h---	h---+QE---SP	h---h---h	h---h---h	h---h---h
telomerase p123	CQPKLFFPATMDIERCYD	REKLSFLKRTTLL-100-RF	REKLSFLKRTTLL-100-RF	REKLSFLKRTTLL-100-RF	REKLSFLKRTTLL-100-RF
Dong (LINE)	KNTRNLHCTYIDYKKKA	FYQTKQTGIFPGLCVSIS	YQTKQTGIFPGLCVSIS	YQTKQTGIFPGLCVSIS	YQTKQTGIFPGLCVSIS
a1 S.c. (groupII)	FDSIPHSVNLIOVLEIYKIN-	QGDSLS?LWFC	QGDSLS?LWFC	QGDSLS?LWFC	QGDSLS?LWFC
HIV-RT	LKKRSVTVIDVGDAYFSVPL	DLRKCFDTISHDLIIXELRYISD-	DLRKCFDTISHDLIIXELRYISD-	DLRKCFDTISHDLIIXELRYISD-	DLRKCFDTISHDLIIXELRYISD-
L8543 . 12 yrs	DKSCTDSDIPRMECMRILDAKDN-	DEKSYCDSDIPRMECMRILDAKDN-	DEKSYCDSDIPRMECMRILDAKDN-	DEKSYCDSDIPRMECMRILDAKDN-	DEKSYCDSDIPRMECMRILDAKDN-
		7-GIRYQYNNVLPCGKGS	7-GIRYQYNNVLPCGKGS	7-GIRYQYNNVLPCGKGS	7-GIRYQYNNVLPCGKGS
		PAIFQSSMTKILPFRKQN	PAIFQSSMTKILPFRKQN	PAIFQSSMTKILPFRKQN	PAIFQSSMTKILPFRKQN
		68-RKYIREDGLFQESSLSA?	68-RKYIREDGLFQESSLSA?	68-RKYIREDGLFQESSLSA?	68-RKYIREDGLFQESSLSA?

FIGURE 18

telomerase p43
human La
Xenopus LaA
Drosophila La
S. c. Lhp1p

LQK**QLEFYFSDANLYNDSFLRKLVLSGEQRVEIETLLM**
ICH**QEYYFGDFNLPRDKFLKEQI**.KLDEGWVPLEIMIK
ICE**QEYYFGDHNLPRDKFLKQQI**.LLDDGWVPLETMIK
ILR**QEYYFGDANLNRDKEFLREQIGRNEDGWVPLSVLVT**
CL**KOMEFYSEFNFPYDRELRTTAEK**.NDGWVPISTIAT

FIGURE 19

FIGURE 20

MEIENNQAQQPKAEKLWWELELEMQENQNDIQVRVKIDDPKQYL
VNVTAAACLQEGSYYQDKDERRYIITKALLEVAESDPEFICQLAVYIRNELYIRTTN
YIVAFCVVHKNTQPFIEKYFNKA VLLPNDLLEVCEFAQVLYIFDATEFKNLYLDRLS
QDIRKELTFRKCLQRCVRSKFSEFNEYQLGKYCTESQRKKTMFRYLSVTNKQKWDQTK
KKRKENLLTKLQAIKESEDKS KRETGDIMNVEDAIAKALKPAVMKKIAKRQNAMKKHMK
APKIPNSTLESKYLTFKDLIKFC HISEPKERVYKILGKKYPKTEEEYKAAFGDSASAP
FNPELAGKRMKJEISKTWENELSAKGNTAEVWDNLISSNQLPYMAMLRNLSNILKAGV
SDTTHSIVNKICEPKAVENSKMFPLQFFSAIEAVNEAVTKGFKA KKRENMLKGQIE
AVKEVVEKTDEEKDMELEQTEEGEFVKVNEGIGKQYINSIELAIKIAVNKNLDEIKG
HTAIFSDVSGSMSTSMSGAKKYGSVRTCLECALVLGLMVQRCEKSSFYIFSSPSSQ
CNKCYLEVDLPGDELPSMQKLLQEKGKLGGGTDFPYECIDEWTKNKTHVDNIVLSD
MMIAEGYSDINVRGSSI VNSIKKYKDEVNPNIKIFA VDLEGYGKCLNLGDEFNENNYI
KIFGMSDSILKFISAKQGGANMVEVIKNFALQKJGQK

FIGURE 21

1 tcaatactat taattaataa ataaaaaaaaa gcaaactaca aagaaaatgt caaggcgtaa
61 ctaaaaaaaaaag ccataggctc ctataggca tgaaacaaat ctgatttg tattacaaaa
121 tctagaagt tacaaaagcc agattgagca ttataagacc tagtagtaat agatcaaaga
181 ggaggatctc aagctttaa agtcaaaaaa ttaagattag gatggaaact ctggcaacga
241 ttagatgtat gaagaaaaca actcaaataa ataataagaa ttataagga gatgtcaatta
301 gattaagtag caagttaat tgataaaaaa agttggctc aaggtagaga aagattgaa
361 ttgaacgaa gatgaaaaca aaaagaatgg acttctgaa tagcaagtga aagaagagta
421 attaagaacg attactgaag aataggtaa gtattaaaat ttagtattta acatggacta
481 ccagttatgat ttaaatgaga gtggggcca tagaagacac agaagagaaa cagattatga
541 tactgaaaaa tggttgaaa tatctcatga ccaaaaaat tatgtatcaa ttacgc当地
601 ctAAAAGACA tcataattgtt ggtggctaa agattattt aataaaaaca attatgatca
661 tcttaatgtt agcattaaca gactagaaac tgaagccgaa ttctatgc当地 ttgatgatt
721 ttcacAAACA atcaaactta ctaataattc ttactagact gttacatag acgttaattt
781 tgataataat ctctgtatac tcgcattgtc tagatTTTA ttatctactg aaagattca
841 tattttgaat ataagatctt ttatataaag aatattat aattttgaga aaatttggta
901 gctacttgc当地 actatctcg cagtgtctt ttctatcg cacttacaag gcatttc当地
961 acaaggccc当地 tgcgaagcgt tctaataattt agttaactcc tcatacataa ttacgc当地
1021 agatgtcaat ttagatgtat actctttctc tacagactta aaatttagtt acactaaca
1081 agtccaagat tatttaagt tcttataaga attccctcg tttactcatg taagcttaga
1141 ggctatcccc gttatgtcta ctaacgc当地 agagaaccctc aatgttttac ttAAAAGGT
1201 caagcatgtc aatcttaattt tagtttctat ccctacctaa ttcaattttt atttctactt
1261 tggtaatttta taacatttga aatttagatgt tggatttagaa ccaaatttt tgacaaaaca
1321 aaagcttgc当地 aatctacttt tgatgtataaa ataatcaaaa aatcttaat tttaagatt
1381 aaacttttac acctacgttgc当地 ttaagaaaac ctccagaaaa cagatattaa aacaagctac
1441 aacaatcaaa aatctcaaaa acaataaaaaa tcaagaagaa actccgtaaa ctaaagatga
1501 aactccaaagc gaaaggccaa gtggatgtaa atttttttgc当地 catctttcg当地 aatcaaccga
1561 gcttgaagat ttctatcg tttatgttgc当地 tacccaaagaa attttatgata gcttgc当地
1621 acttttgattt agatcaacaa attttaaagaa gttcaatttta agtttacaaaat atgaaatgg
1681 aaagatgttgc当地 atggatcat tcatagatct taagaatattt tttatgttgc当地
1741 taaaatgttgc当地 tttatgttgc当地 tttatgttgc当地 aatcttttgc当地
1801 taaaatgttgc当地 aatcttttgc当地 aatcttttgc当地 aatcttttgc当地
1861 gttatctttt aatgttgc当地 aatcttttgc当地 aatcttttgc当地 aatcttttgc当地
1921 ttccatcaat gaaagcttgc当地 aatgttgc当地 aatcttttgc当地 ttatgttgc当地
1981 aaatttacaa aatgttgc当地 tttatgttgc当地 cccaaatataa tttagaaaaaa
2041 tccttcaat aagcccaatc ttatgttgc当地 caagcaattt gaataatttga aatgttgc当地
2101 aatgttgc当地 atcaacttgc当地 ttatgttgc当地 gcatataattt cagaatttgc当地
2161 agaaaagaat aaaaaataaa aagcatttgc当地 ttgttgc当地 tttatgttgc当地
2221 tcttgc当地 tttatgttgc当地 aatgttgc当地 ttatgttgc当地
2281 cattatgttgc当地 caatgttgc当地 aatgttgc当地 gagtttgc当地 cataatgttgc当地
2341 ccacaagcaa aatgttgc当地 atgttgc当地 aatgttgc当地 atcaatgttgc当地
2401 ctttgc当地 atgttgc当地 accaaaacac tttatgttgc当地 gatgttgc当地 aatgttgc当地
2461 agaatcttgc当地 tttatgttgc当地 agtcatcaat ttatgttgc当地 ttatgttgc当地
2521 ctttgc当地 aatgttgc当地 aatgttgc当地 aatgttgc当地 tttatgttgc当地
2581 agggttttgc当地 gtttgc当地 aatgttgc当地 ctttgc当地 aatgttgc当地
2641 ctttgc当地 tttatgttgc当地 gtttgc当地 aatgttgc当地 tttatgttgc当地
2701 tttatgttgc当地 tttatgttgc当地 aatgttgc当地 tttatgttgc当地
2761 atatatttttgc当地 gtttgc当地 aatgttgc当地 tttatgttgc当地
2821 aaaaaatcg

FIGURE 22

MSRRNQKKPQAPIGNETNLDFVLQNLEVYKSQIEHYKTQQQQIK
EEDLKLLKFKNQDQDGNSGNDDEENNSNKQQELLRRVNQIKQQVQLIKVGSKVEK
DLNLNEDEKKNGLSEQQVKEEQLRTITEEQVKYQNLVFNMODYQLDLNESGGHRRHRR
ETDYDTEKWFEISHDQKNYVSIYANQKTSYCWWLKDYFNKNYYDHLNVSVIRLETEAE
FYAFDDFSQTIKLTNNSYQTVNIDVNFDNNLCILALLRFLLSLERFNLNIRSSYTRN
QYNFEKIGELLETIFAVVFSHRHLQGIHLQVPCEAFQYLVNSSSQISVKDSQLQVYSF
STDLKLVDTNKVQDYFKFLQEFPRLTHVSQQAIPVSATNAVENLNVLLKKVKHANLNL
VSIPTQFNFDFYFVNQLQHLKLEFGLEPNILTQKLENLLSIKQSKNLKFLRLNFYTY
VAQETSRKQILKQATTIKNLKNNKNQEETPETKDETPSESTSGMKFFDHLSELTELED
FSVNLQATQEIYDSLHKLLIRSTNLKKFKLSYKYEMEKSMDTFIDLKNIYETLNNLK
RC SVNISNP HGNI SYEL TNK D STF YKF KLT LNQ ELQ HAK YT F K Q NEF Q F N V K S A K I E
SSSLESLEDIDS LCKS IAS CKNLQ NVN II AS L L Y P N N I Q K N P F N K P N L L F F K Q F E Q L K
NLEN VS IN CILDQH IL NS I SEF LEK N K K IKA FIL K R Y Y LL Q Y Y LD Y T K L F K T L Q Q L P E
LNQVYINQQLEELTVSEVHKQVWENHKQKAFYEPLCEF IK ESS Q TL Q L ID FD Q N TVSD
DSIKKILESISESKYHHYLRLNPSQSSLIKSEN E EI Q ELL KAC DEKG VL VKAYYKFP
LCLPTGTYYDYN SDR W

FIGURE 23

MKILFEFIQDKLDIDLQTNSTYKENLKGHFNGLDEILTCFAL
PNSRKIALPCLPGDLSHKAVIDHCIYLLTGELEYNNVLTFGYKIARNEDVNNSLFCHS
ANVNVTLLKGAAWKMFHSLVGTYAFVDLLNYTVIQFNGQFFTQIVGNRCNEPHLPPK
WVQRSSSSATAAQIKQLTEPVTKQFLHKLNINSSSFPYSKILPSSSIKKLTDLR
EAIFPTNLVKIPQRLKVRJNLTQKLLKRHKRLNYVSILNSICPLEGTVLDLSHLSR
QSPKERVLKFIIVILQKLLPQEMFGSKKNKGKIIKNLNLLSLPLNGYLPFDSSLKKL
RLKDFRWLFISDIWFTKHNFENLNQLAICFISWLFRQLIPKJIQTFFYCTEISSTVTI
VYFRHDTWNKLITPFIVEYFKTYLVENNVCRNHNSYTLSNFNHSKMRIIPKKSNNEFR
IIAIPCRGADEEEFTIYKENHKNAIQPTQKILEYLRNKRPTSFTKJYSPTQIADRIKE
FKQRLKKFNNVLPELYFMKFDVKSCYDSIPRMECMRILKDALKNENGFFVRSQYFFN
TNTGVLKLFNVVNASRVPKPYELYIDNVRTVHLSNQDVINVVEMEIFKTALWVEDKCY
IREDGLFQGSSLSAPIVDLVYDDLLEFYSEFKASPSQDTLILKLAADDFLIISTDQQQV
INIKKLAMGGFQKYNAKANRDKILA VSSQSDDDTVIQFCAMHIFVKELEVWKHSSTMN
NFHRSKSSKGIFRSLIALFNTRISYKTIDTNLNSTNTVLMQIDHVVKNISECYKSAF
KDLSINVQTQNMQFHSFLQRIIEMTVSGCPITKCDPLIEYEVRFTILNGFLESLSSNTS
KFKDNIILLRKEIQHLQAYIYIYIHVN



FIGURE 24

Oxytricha
Euplotes

LCVSYILSSFYANLEENALQFLRKESMDPEKPETNLLMRLT
LCVSSILSSFYATLEESSLGFLRDESMNPENPNVNLLMRLT

FIGURE 25

human
tez1
EST2
p123

Motif 0
AKFLHWLMSVYVVELLRSPFYVTETTPQKNR
ISPIEVLVLGKRSNAKMCLSDFEKRKQIIFAYWLYNSPIIPIQSFFYITESSDLRNR
LKDFRWLFISD---IWFTKHNFENLNQLAICFISWLPRQLIPKIIQTFFYCTEISSTVT-
TREISWMQVET-SAKHYYFDHEN-IYVLWKLLRWIFEDLVVSLIRCFYVTEQQKSYSK
.....***

human
tez1
EST2
p123

Motif 1
LFFYRKSVWSKLQSIGIROHLKRVQLRDVSEAEVRQHREARPALLTSRLRPIPKP--DGL
TVYFRKDIWKLLCRPPI-TSMKMEAFEKINENNVRMDTQK-TTLPPAVIRLLPKK--NTP
IVYFRHDTWNKLITPFIVEYFKTYLVENNVCRNHNSYTLS--RFNHSKMRIIPKKSNNEF
TYYYRKNIWDVIMKMSI-ADLKKETLAEVQEKEVBEWKKS-LGFAPGKLRLIPKK--TTF
...*

human
tez1
EST2
p123

Motif 2
RPIVNMDYVVGARTFRREKRAERLTSRVKALP-8VLYERA
RLITN-LRKRPlikmgnskkmlvstnqtlrpvasilkhlneessgipfnlevymklltf
RIIAIPCRGADEEEFTIYKENHKNAIQPTQKILEYLRNKRPTSFTKIYSPTQIADRIKEF
RPIMTFNKKIVNSDRKTTKLTNTKLLNSHMLKTLKN-RMFKDPPGPAVFNYDDVMKKY
**

human
tez1
EST2
p123

Motif 3 (A)
KKDLLKHRMFGR-KKYFVRIDIKSCYDRIKQDLMFRIVKR-KLKDPEPVIRKYATIHATS
KQRLLKKFNNVLPELYFMKFDVKSCYDSIPRMEMCMLKID-ALKNENGFFVRSQYFFNTN
KEFVCKWKQVGQPKLFFATMDIEKCYDSVNREKLSTFLKTTKLLSSDFWIMTAQILKRKN
...*

FIGURE 26

ATTTATACTCATGAAAATCTTATTCGAGTTCATTCAAGACAAGCTGACATTGATCTACA
GACCAACAGTACTTACAAAGAAAATTAAAATGTGGTCATTCAATGGCCTCGATGAAAT
TCTAACTACGTGTTCGCACTACCAAATTCAAGAAAATAGCATTACCATGCCCTCCTGG
TGACTTAAGCCACAAAGCAGTCATTGATCACTGCATCATTACCTGTTGACGGGCGAATT
ATACAACAACGTAACACATTGGCTATAAAATAGCTAGAAATGAAGATGTCAACAAATAG
TCTTTTTGCCATTCTGCAAATGTTAACGTTACGTTACTGAAAGGGCGTGTGGAAAAT
GTTCCACAGTTGGTCGGTACATACGCATTGTTGATTATTGATCAATTACAGTAAT
TCAATTAAATGGGAGTTTCACTCAAATCGTGGGTAACAGATGTAACGAAACCTCATCT
GCCGCCAAATGGGTCCAACGATCATCCTCATCATCCGCAACTGCTGCGCAAATCAAACA
ACTTACAGAACCGAGTACAATAAAACAATTCTACACAAGCTCAATATAAATTCCCTCTTC
TTTTTTCTTATAGCAAGATCCTTCCTTCATCATCATCTATCAAAAAGCTAAGTACCT
GAGAGAAGCTATTTTCCCACAAATTGGTTAAAATTCTCAGAGACTAAAGGTACGAAT
TAATTGACGCTGCAAAGCTATTAAAGAGACATAAGCGTTGAATTACGTTCTATT
GAATAGTATTTGCCACCATTGGAAGGGACCGTATTGAACTTGTGCAATTGAGTAGGCA
ATCACCAAAGGAACGAGTCTGAAATTATCATTGTTATTACAGAAGTTATTACCCCCA
AGAAATGTTGGCTCAAAGAAAATAAGGAAAATTATCAAGAATCTAAATCTTTATT
AAGTTACCCCTAAATGGCTATTACCATTTGATAGTTGTTGAAAAGTTAAGATTAAA
GGATTTCGGTGGTTGTCATTCTGATATTGGTCACCAAGCACAATTGAAAACCTT
GAATCAATTGGCATTGTTCATTCCTGGCTATTAGACAACTAATTCCAAAATTAT
ACAGACTTTTTACTGCACCGAAATATCTCTACAGTACAATTGTTACTTAGACA
TGATACTTGAATAAAACTTATCACCCCTTTATCGTAGAATATTAAAGACGTACTTAGT
CGAAAACAACGTATGTAGAAACCATAATAGTTACACGTTGCAATTCAATCATAGCAA
AATGAGGATTATACCAAAAAAAAGTAATAATGAGTTCAAGGATTATTGCCATCCATGCAG
AGGGGCAGACGAAGAAGAATTACAATTATAAGGAGAACACAAAATGCTATCCAGCC
CACTAAAAAATTAGAATACCTAAGAAACAAAAGGCCGACTAGTTACTAAATATA
TTCTCCAACGCAAATAGCTGACCGTATCAAAGAATTAAAGCAGAGACTTTAAAGAAATT
TAATAATGTCTTACAGAGCTTATTTCATGAAATTGATGTCATCTGCTATGATTC
CATACCAAGGATGGAATGTATGAGGATACTCAAGGATGCGCTAAAAAATGGGTT
TTTCGTTAGATCTCAATATTCTCAATACCAATACAGGTGTATTGAAGTTATTAAATGT
TGTTAACGCTAGCAGACTACCAAAACCTTATGAGCTATAAGATAATGTGAGGACGGT
TCATTATCAAATCAGGATGTTAAACGTTGAGAGATGAAATATTAAACAGCTT
GTGGGTTGAAGATAAGTGTACATTAGAGAACAGTGGTCTTTCAAGGCTCTAGTTATC
TGCTCCGATCGTGTATTGGTGTATGACGATCTCTGGAGTTTATAGCGAGTTAAAGC
CAGTCCTAGCCAGGACACATTAATTAAAAGCTGGCTGACGATTCTTATAATATCAC
AGACCAACAGCAAGTGTCAATATCAAAAAGCTTGCATGGCGGATTCAAAAATATAA
TGCAGAAAGCCAATAGAGACAAAATTAGCCGTAAGCTCCAACTCAGATGATGATACGGT
TATTCAATTGTGCAATGCACATATTGTTAAAGAATTGGAAGTTGGAAACACATTCAAG
CACAATGAATAATTCCATATCGTTGAAATCTAGTAAAGGGATATTGAAAGTTAAT
AGCGCTGTTAACACTAGAATCTCTTATAAAACAATTGACACAAATTAAATCAACAAA
CACCGTTCTCATGCAAATTGATCATGTTGAAAGAACATTGGAATGTTATAATCTGC
TTTAAGGATCTATCAATTATGTTACGAAAATATGCAATTCTGTTCTTAAACGAGTA
CATCATTGAAATGACAGTCAGCGGTTGTCATTACGAAATGTGATCCTTAAATCGAGTA
TGAGGTACGATTCAACATTGAATGGATTGGAAAGCCTATCTCAAACACATCAA
ATTAAAGATAATATCATTCTTGGAGAAAGGAAATTCAACACTTGCAAGC

FIGURE 27

AKFLHWLMSVYVVELRSFFYVTETTFQKNRLFFYRKSWSKLQSIGIRQHLKR
VQLRDVSEAEVRQHREARPALLTSRLRFIPKPDGLRPIVNMDYVVGARTFRREKR
AERLTSRVKALFSVLNYERA

FIGURE 28

GCCAAGTTCTGCACGGCTGATGAGTGTGTACGTCGTCAGCTGCTCAGGTC
TTTCTTTATGTCACGGAGACCACGTTCAAAAGAACAGGCTCTTTCTACC
GGAAGAGTGTCTGGAGCAAGTTGCAAAGCATTGGAATCAGACAGCACTGAA
GAGGGTGCAGCTGCCGGACGTGTCGGAAGCAGAGGTCAAGCAGCATCGGGA
AGCCAGGCCGCCCCTGCTGACGTCCAGACTCCGCTTCATCCCCAAGCCTGACG
GGCTGCGGCCGATTGTGAACATGGACTACGTCGTGGGAGCCAGAACGTTCCG
CAGAGAAAAGAGGGCCGAGCGTCTCACCTCGAGGGTGAAGGCACTGTTCAGC
GTGCTCAACTACGAGCGGGCGCG

FIGURE 29

MTEHHTPKSRJLRFLENQYVYLCTLNDYVQLVLRGSPASSYSNICERLRSDVQTSFSIFLHSTVVGF
DSKPDEGVQFSSPKCSQSELIANVVQMFDESERRNLLMKGFSMNHEDFRAMHNGVQNDLV
STFPNYLISILESKNWQLLEIIGSDAMHYLLSKGSIFEALPNDNYLQISGIPLFKNNVFEETVSKRK
RTIETSITQNKSARKEVSNSISRSIFSIFYRSSYKKFKQDLYFNLHSICDRNTVHMWLQWIFPRQFG
LINAFAQVKQLHKVIPLVSQSTVVPKRLLKVYPLIEQTAKRLHRISLSKVYNHYCPYIDTHDDEKILS
YSLKPNQVFALRSILVRVFPKLIWGNQRIFEIILKDLETFLKLSRYESFSLHYLMSNIKISEIEWLVL
GKRSNAKMCLSDFEKRKQIFAEIFYWLYNSFIIPILQSFFYITESSDLRNRTVYFRKDIWKLLCRPFIT
SMKMEAFEKINENNVRMDTQKTTLPPAVIRLLPKKNTFRLITNLRKRFLIKMGSNKKMLVSTNQT
LRPVASILKHLINEESSGIPFNLEVYMKLLTFKKDLLKHRMFGRKKYFVRIDIKSCYDRIKQDLMFR
IVKKKLKDPEFVIRKYATIHATSDRATKNFVSEAFSYFDMVPFEKVVQLLSMKTSDTLFVDFVDY
WTKSSSEIFKMLKEHLSGHIVKIGNSQYLQKVGPQGSILSSFLCHFYMEDLIDEYLSFTKKKGSVL
LRVVDDFLFITVNKKDAKKFLNLSLRGFKEKFHNFSTSLEKTVINFENSNGIINTFFNESKKRMPFFG
FSVNMRSLDTLLACPKIDEALFNSTSVELTKHMGKSFFYKILRSSLASFAQVFIDITHNSKFNSCCNI
YRLGYSMCMRAQAYLKRMKDIFIPQRMFITDLLNVIGRKIWKLLAEILGYTSRRFLSSAEVKWLFC
LGMRDGLKPSFYHPCFEQLIYQFQSLTDLIKPLRPVLRQVLFLHRRJAD

FIGURE 30.

FIGURE 30 (cont.)

ATGgtacgtgcggctcgagactcagcaatattgacacatcagGCTTTTGTCTGGAATGAGAGATGGTTGAAACCCTCTT
TCAAATATCATCCATGCTTCGACAGCAGCTAATATACCAATTTCAGTCATTGACTGATCTTATCAAGCCGC
TAAGACCAGTTTGCGACAGGTGTTATTTTACATAGAAGAATAGCTGATTAAtgcattttcaatttattatacatccctt
tattactggtgtcttaaaacaatattactaagtatagctgaccccaaagcaagcatactataggattctagtaaagtaaaattaatcicgtaatttagtttgattgacttgtctt
atccctatacttttaagaagattgacagtggttgcactactgcccacatgcccattaaacgggagtggtaaacattaaagaataatcatgaggctaattcccttcatttag
aataaggaaagtggttttctataatgaataatgcccgactaatgcaaaaagacaaggattatctctaaacaagggggatlaagcatatccgaaggaaaagagagataatat
accagggttgtgaagaagcaaggataatttgaacaagcttctgcagatgacaggctaaattttggtgaccgaattttggtaaaagccccaggtaatccatgggtggccg
gccctgcactgagacgaaaagaactaaggatagtttgaataactaatacgcttataatgtcttatataagggtttttccctgacttcaatttgcatggtgaaaagaaaata
gtgttaagccattttggattccgaaaatagccaaattttctggttctcaaaagcggaaagtctaaagaacttattgaagctttaggcttcaaaaactccctgatttaaaggag
gaatctccaccgatgaggaaatggatagcttacagctgcgtgaggagaagcctaatttttgcaaaaaaagaaaatatcattgggagacatcttgaatcagatgcgga
gagtatctccagcggatccctgtatcaataacttcttctgaaatgtatggtctactgtcgctcactctcgttagctcactgcagttaaagtgaccaaaggtaacc

FIGURE 31

EST2 pep	FFYCTEISST	VTIVYFRHDT	WN---	KLIT	P----	FIVE	YFK-TYLVEN	40
Euplotes pep	FFYVTEQQKS	YSKTYYYRKN	IWDVI-MKMS	IAD---	LKK	ETLA--EVQE		43
Trans of tetrahymen	-----KHKE	GSQIFYYRK	IWKLVSKLT	VKVRIQFSEK	NKQMKNNFYQ			44
Consensus	FFY.TE..K.	.S..YYYRK.	IW...-KL..-F..KV..			50
EST2 pep	NVCRNHNSY-	-----	TLSNFNHSKM	RUTIPKKSSNE	FRUITAIPCRG			79
Euplotes pep	KEVEEWKKSL	-----	--GFAPGK	RUTIPKKIT	FRUITMTFNKK			78
Trans of tetrahymen	KIQLEENLE	KVEEKLIPE	SFQKYPQCK	RUTIPKKGS	FRUITMTFLRK			92
Consensus	K...E.....	-----	...F...GK	RUTIPKKI	FRUITMTF.RK			100
EST2 pep	ADEEEFTIYK	ENHKNAIQPDI	OKILEYERKN	RPTSFTKIYS	PTQLADRIKE			129
Euplotes pep	IVNSDRKTTK	LTTNTKLNS	HLMLKTKLKN	-----	RMFK	-DPFGFAVFN		120
Trans of tetrahymen	DKQKNIK	---LNLNQILMDS	OLVFRNKD	-----	YL-G	-QKIGYSVFD		130
ConsensusK..K	LN.N..L..S	:QL.L..EKN	-----	-..IG..VF.		150
EST2 pep	FKQRLLKKFN	NVL-----	PELEYFMKFD	VKSCYD				157
Euplotes pep	YD-DVMKKYE	EFVCKWKQVG	QPKUFFATMD	IEKCYD				155
Trans of tetrahymen	NK-QISEKFA	QFIEKWKNKG	RPOHYVTL	-----				158
Consensus	.K...KKF..	.F..KWK..G	PELEYF.T.D	...CYD				186

FIGURE 32

S-1: FFY VTE TTF QKN RLF FYR KSV WSK

S-2: RQH LKR VQL RDV SEA EVR QHR EA

S-3: ART FRR EKR AER LTS RVK ALF SVL NYE

A-1: AKF LHW LMS VYV VEL LRS FFY VTE TTF Q

A-2: LFF YRK SVW SKL QSI GIR QHL KRV QLR DVS

A-3: PAL LTS RLR FIP KPD GLR PIV NMD YVV

FIGURE 33

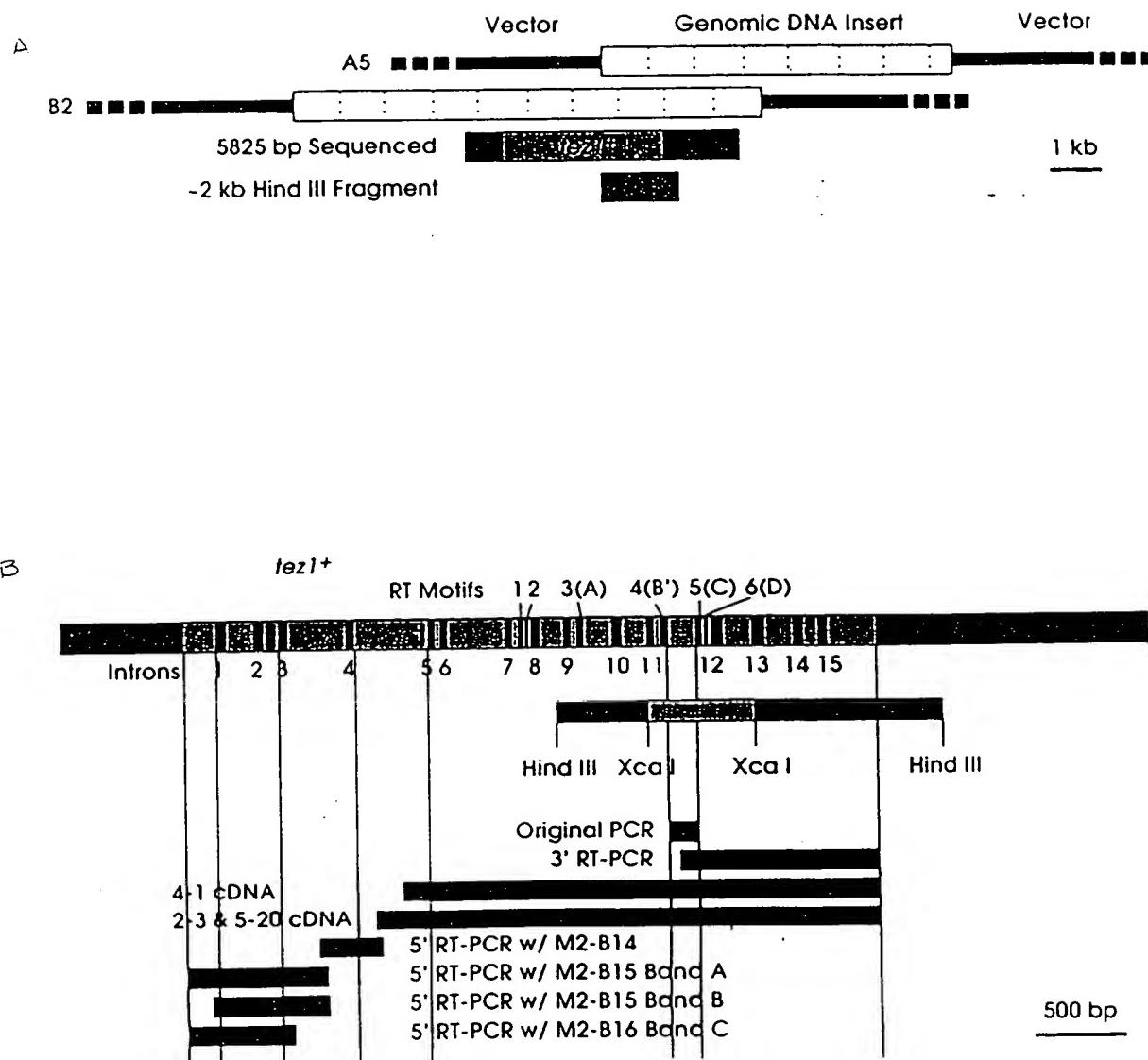


FIGURE 34

Poly 4

 t t c
 t a a g c c t c g
5' - cag acc aaa gga att cca taa gg -3'
 Q T K G I P Q G

4 (B')

5 (C')

 D D Y L L I T
3' - ctg ctg atg gag gag tag tgg -5'
 a a a a a a a a a
 t t t t
 c c
Poly 1

FIGURE 35

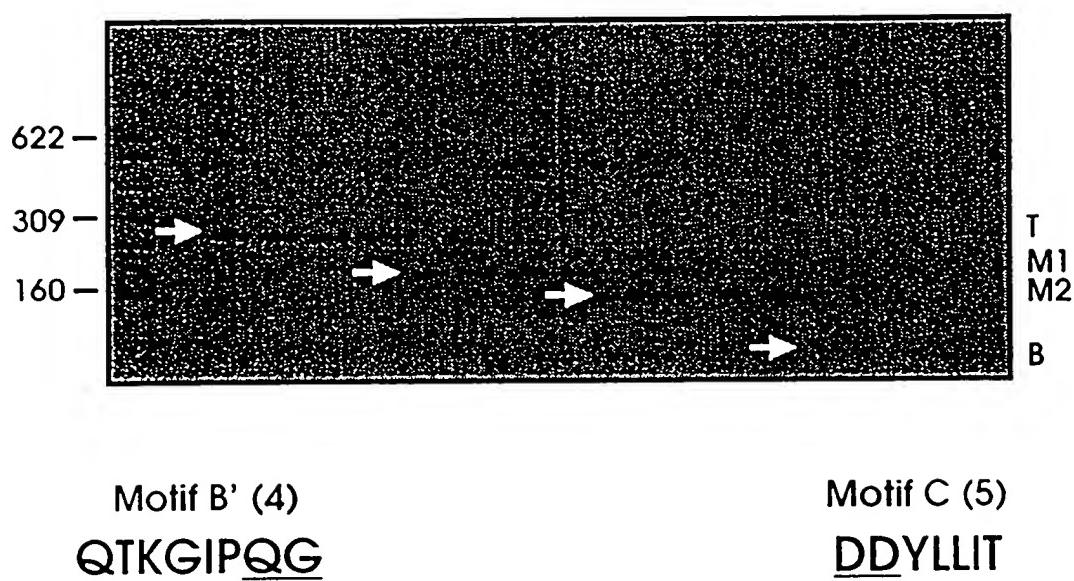


FIGURE 36

PCR Product M2 showed Reasonable Match with Other Telomerase Proteins

ot	LCVSYILSSFYANLEENALQFLRKESMDPEKPETNLLMRLT
Ea_p123	KGIPQGLCVSSILSSFYYATLEESSLGFLRDESMNPENPNVNLLMRLTDDYLLIT
Sp_M2	SILSSFLCHFYMEDLIDEYLSFTKKK-----GSVLLRVV
Sc_p103	DGLFQGSSLSAPIVDLVYDDLLEFYSEFKASPS-----QDTLILKLADDFLIIS

Q K V G I P Q G
qaa aaa gtt ggt atc cct cag gg..... <---Actual Genomic Sequence.

Poly 4

t t c
a a g c c t c g
Cag acc aaa gga att cca taa gg ---->

ag acc aaa gga att cca tca ggC TCA ATT CTG TCA TCT TTT TTG TGT CAT TTC TAT ATG
tc tgg ttt cct taa ggt agt ccG AGT TAA GAC AGT AGA AAA AAC ACA GTA AAG ATA TAC

K G I P S G S I L S S F L C H F Y M

GAA GAT TTG ATT GAT GAA TAC CTA TCG TTT ACG AAA AAG AAA GGA TCA GTG TTG TTA CGA
CTT CTA AAC TAA CTA CTT ATG GAT AGC AAA TGC TTT TTC TTT CCT AGT CAC AAC AAT GCT

E D L I D E Y L S F T K K G S V L L R

GTA GTC gac gac tac ctc ctc atc acc
CAT CAG ctg ctg atg gag gag tag tgg

V V D D Y L L I T

<---- ctg ctg atg gag gag tag tgg
a a a a a a a a
t t t t
c c

Poly 1

.....gac gat ttc ctc ttt ata aca..... <---Actual Genomic Sequence.

D D F L F I T

FIGURE 37

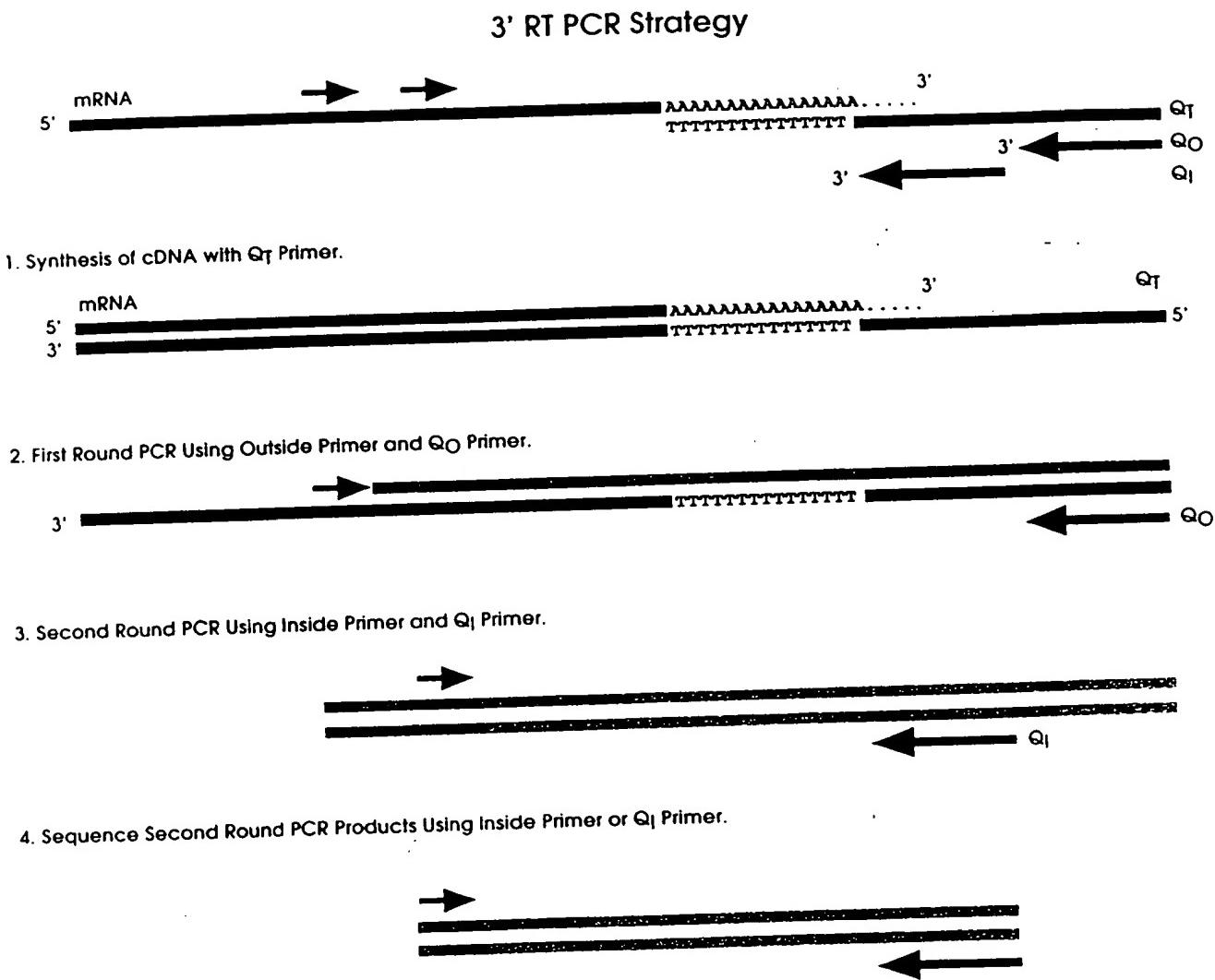


FIGURE 38

A

-Size Selected Libraries from P. Nurese

3 ~ 4 kb
5 ~ 6 kb
7 ~ 8 kb
11 ~ 12 kb

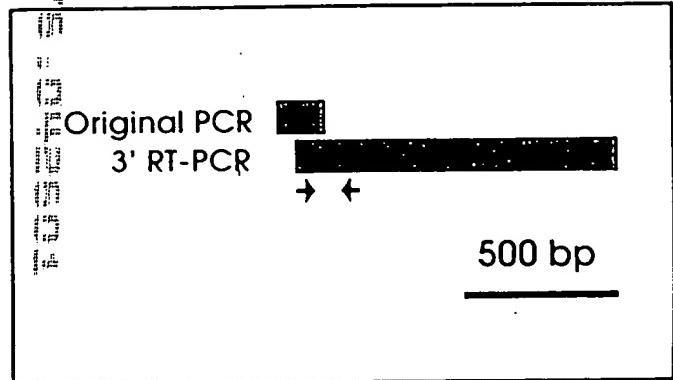
-Libraries from J.A. Wise

Sau 3a Partial Digest
Hind III Partial Digest

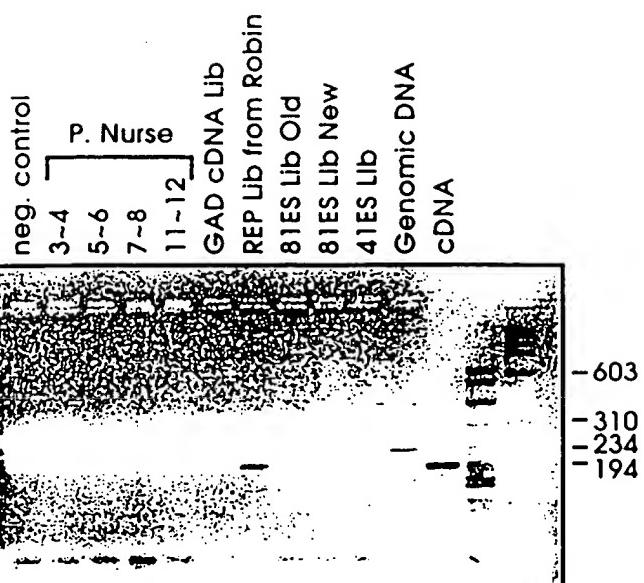
cDNA Libraries

GAD (Gal Activation Domain) Library
REP Library from R. Allshire
REP81ES Library (old)
REP81ES Library (new)
REP41ES Library

B



C



D

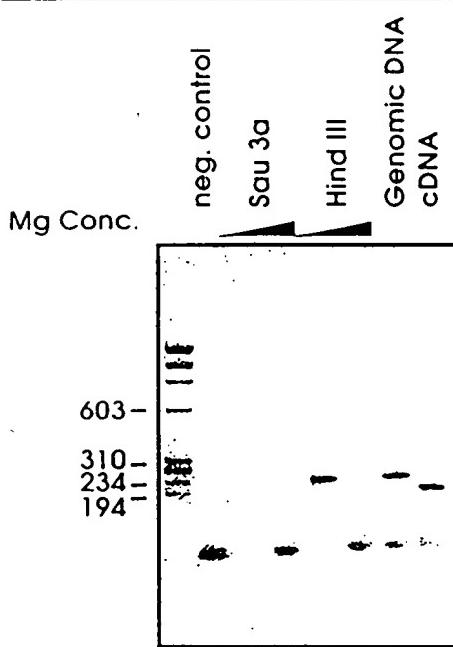


FIGURE 39

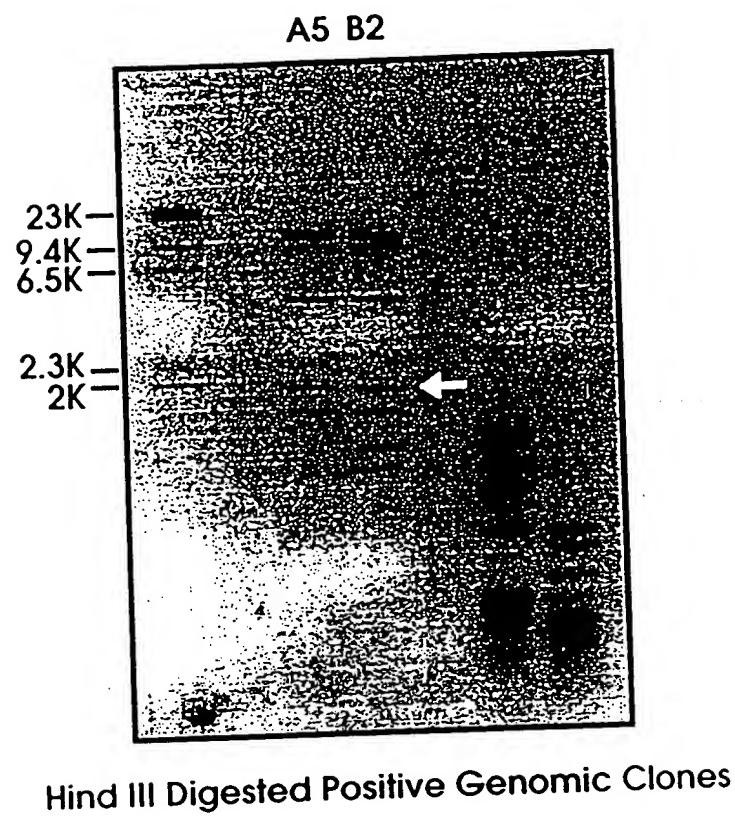
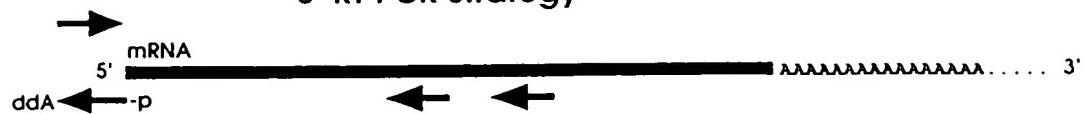
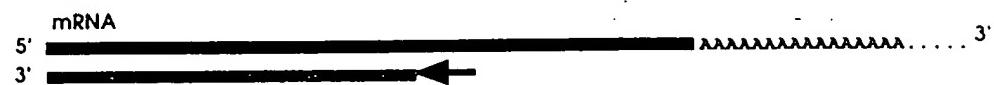


FIGURE 40

5' RT PCR Strategy



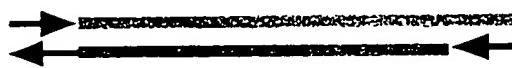
1. Synthesis of cDNA with Specific Downstream Primer.



2. Ligate Oligo with 5'-P and blocked 3' to cDNA using T4 RNA Ligase.



3. First Round PCR



4. Second Round PCR

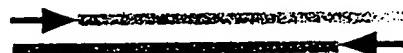


FIGURE 41

Alignment of RT Domains from Telomerase Catalytic Subunits.

Motif 0

S.p. Tez1p (429). WLYNSFIIPILQSFFYITESSDLRNRTVYFRKDIW ... (35) ...
S.c. Est2p (366). WLFRQLIPKIIQTFYCTEISSTVT-IVYFRHDTW ... (35) ...
E.a. p123 (441). WIFEDLVVSLIRCFYVTEQQKSYSKTYYYRKNIW ... (35) ...
* * * * *

Motif 1 Motif 2 K
p hh h K hR h R

S.p. Tez1p AVIRLLPKK--NTFRLITN-LRKRF ... (61) ...
S.c. Est2p SKMRIIPKKSNNEFRIIAIPCGRGAD ... (62) ...
E.a. p123 GKLRLIPKK--TTFRPIMTFNKKIV ... (61) ...
* * * * *

Motif 3(A) AF
h hDh GY h

S.p. Tez1p KKVFVRIDIKSCYDRIKQDLMFRIVK ... (89) ...
S.c. Est2p ELYFMKFDVKSCYDSIIPRMECMRILK ... (75) ...
E.a. p123 KLFFATMDIEKCYDSVNREKLSTFLK ... (107) ...
* * * * *

Motif 4(B')

hPQG pP hh h

S.p. Tez1p YLQKVGIPQGSILSSFLCHFYMEDLIDEYLSF ... (6) ...
S.c. Est2p YIREDGLFQGSSLSSAPIVDLVYDDLLEFYSEF ... (8) ...
E.a. p123 YKQTKGIPQGLCVSSILSSFYATLEESSLGF ... (14) ...
* * * * *

Y Motif 5(C)

h F DDhhh

Motif 6(D)

Gh h cK h

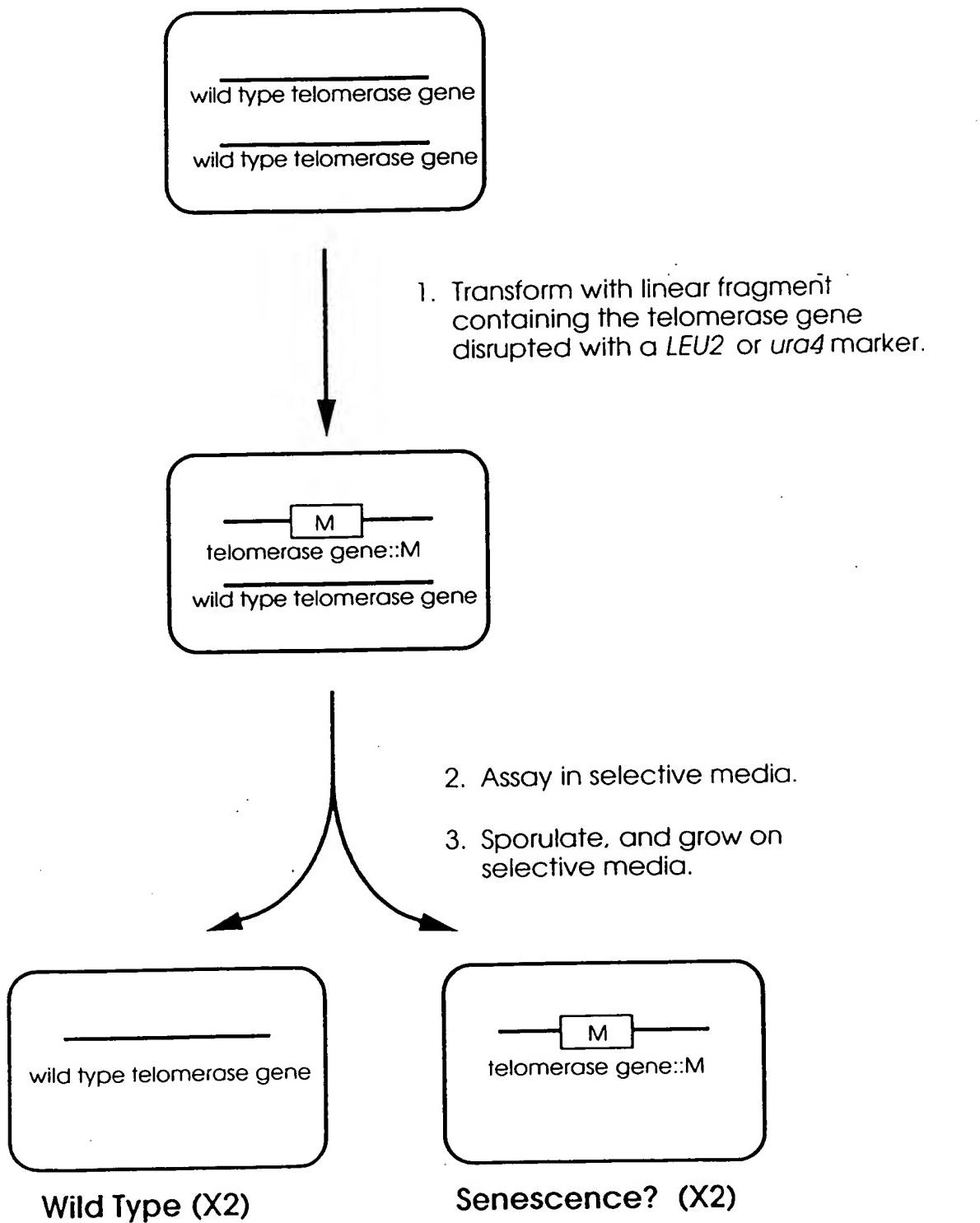
S.p. Tez1p VLLRVVDDFLFITVNKKDAKKFLNLRLGFEKHNFSSTSLEKTVINFENS . (205)
S.c. Est2p LILKLADDFLIISTDQQQVINIKKLAMGGFQKYNAKANRDKILAVSSQS . (173)
E.a. p123 LLMPLTDDYLLITTQENNAVLFIEKLINVSRENGFKFNMKLQTSFPLS . (209)
* * * *

FIGURE 42

A

FIGURE 43

Disruption strategy for the putative telomerase genes.



(These cells will show a senescence phenotype if the disrupted gene encodes a telomerase subunit.)

FIGURE 44

An Example of Confirmation of *tez1* disruption By PCR

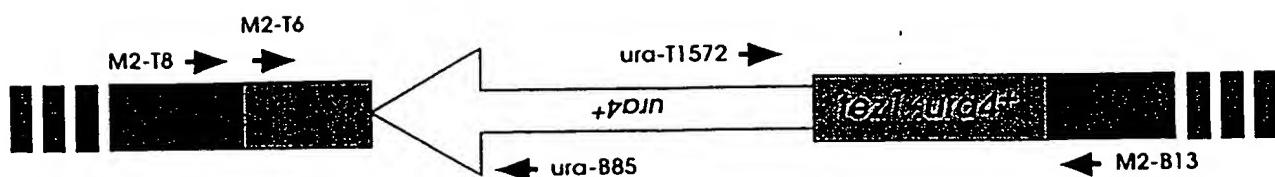
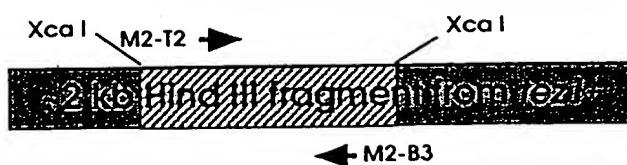
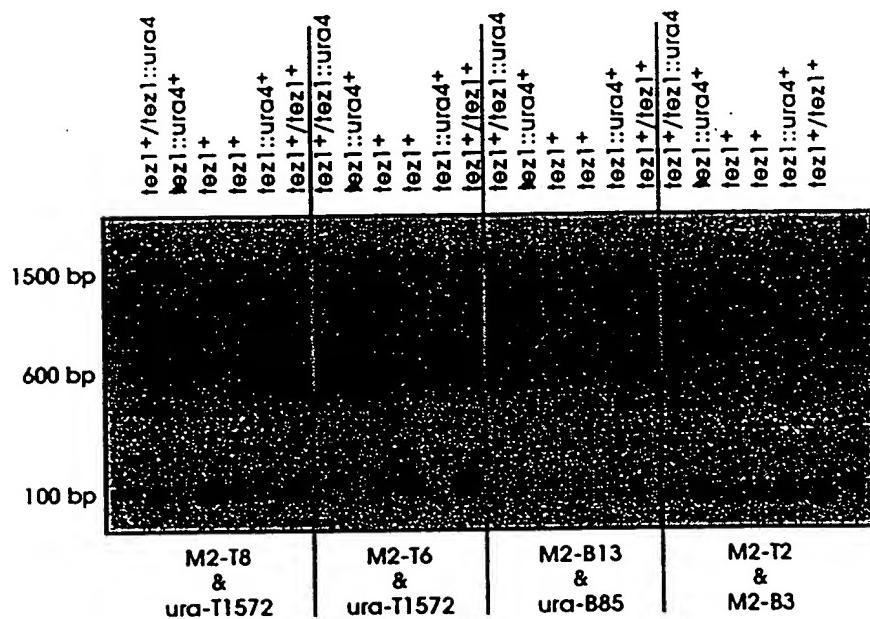


FIGURE 45

Tez1 disruption causes progressive shortening of telomeres in *S. pombe*

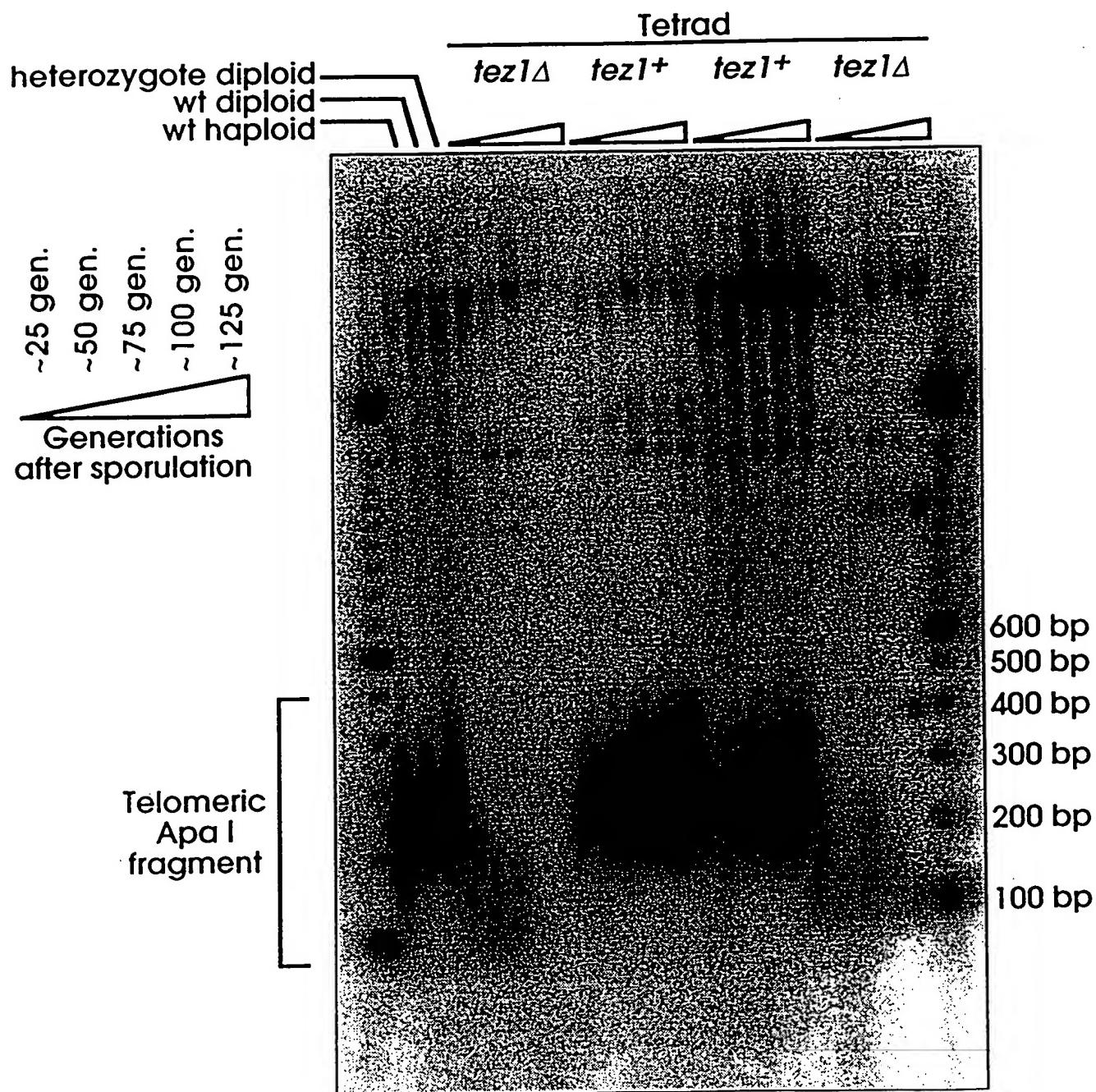


FIGURE 46

1 ggtaccgatttactttccttcataagctaattgcttcctcgAACGCTCCTAAATCTCTGGAAATATTTACAAGA 80
 81 actcaataacaataccaggaaatccaaatgcgggttattagtatcgataatattctatTTATCGGTGTTA 160
 161 ccaagtataaggacaaaagaacaacttcctccccctaaagactttactttattaatttactttcaaatatTCG 240
 241 ggTCGCTTACTTTAATCGTGGTACTGTTAGCTGCTACTCTAGCCAACCGCGTGGTCTACCCGTCATTGGATA 320
 321 agctctggagtagctcacagaaatcTTACAAATCTCTGATGAGACTATTAGATTCAATTACAGTCCTGCATATT 400
 401 ttaacatggagcTTACACTTAGATGAGTCACGTCGATGATGGAGTATTGGATCATCCAAACGTTGCCTGAAAAG 480
 481 gttgataattatttgcAAAATCATGTCCTTAGTGGTGGTAATCCCGCAAAGTTTGATGCTGCACACGTCTAGCATG 560
 561 attgagatattcaaaaatttcatccactacaactcTTAACCGCGTTTATTCTATTCTATTCTATTCTATTGTT 640
 641 ccaaataatgttatcatctcgattaggCTTTCGTTTACTCCTGGAATCGTACCTTTACTATTCCCTAAATGAA 720
 721 ataatctaaatttagttcgTTTATAATTGATAGTAGTAGAAAGATTGGTGAATTCTACTCGTGAATGTTATTGTTAAA 800
 801 gatactttgcAAAACATTATTAGCTATCATTATAAAAAAAATCCTATAATTATAATTCAATATTGCGGTC 880
 881 actatttatttAAAACGTTATGATCAGTAGGACACTTGCAATATAGTTATGCTTAATGGTTACTGTAACCTGC 958

 959 ATG ACC GAA CAC CAT ACC CCC AAA AGC AGG ATT CTT CGC TTT CTA GAG AAT CAA TAT GTA 1018
 1 M T E H H T P K S R I L R F L E N Q Y V 20

 1019 TAC CTA TGT ACC TTA AAT GAT TAT GTA CAA CTT GTT TTG AGA GGG TCG CCG GCA AGC TCG 1078
 21 Y L C T L N D Y V Q L V L R G S P A S S 40

 1079 TAT AGC AAT ATA TGC GAA CGC TTG AGA AGC GAT GTA CAA ACG TCC TTT TCT ATT TTT CTT 1138
 41 Y S N I C E R L R S D V Q T S F S I F L 60

 1139 CAT TCG ACT GTA GTC GGC TTC GAC AGT AAG CCA GAT GAA GGT GTT CAA TTT TCT TCT CCA 1198
 61 H S T V V G F D S K P D E G V Q F S S P 80

FIGURE 46 (cont.)

1199 AAA TGC TCA CAG TCA GAG gtatatatattttgtttgatttttctattcggatagctaataatatggcag	1272
81 K C S Q S E	86
1273 CTA ATA GCG AAT GTT GTA AAA CAG ATG TTC GAT GAA AGT TTT GAG CGT CGA AGG AAT CTA	1332
87 L I A N V V K Q M F D E S F E R R R N L	106
1333 CTG ATG AAA GGG TTT TCC ATG gtaaggattctaattgtgaaatattacacctgcaattactgtttcaaagaga	1405
107 L M K G F S M	113
1406 ttgtatccaaccgataaaag AAT CAT GAA GAT TTT CGA GCC ATG CAT GTA AAC GGA GTA CAA AAT	1469
114 N H E D F R A M H V N G V Q N	128
1470 GAT CTC GTT TCT ACT TTT CCT AAT TAC CTT ATA TCT ATA CTT GAG TCA AAA AAT TGG CAA	1529
129 D L V S T F P N Y L I S I L E S K N W Q	148
1530 CTT TTG TTA GAA AT gtaaaataccggtaagatgtgcgcacttgaacaagactgacaagtata T ATC GGC	1601
149 L L L E I	I G 155
1602 AGT GAT GCC ATG CAT TAC TTA TTA TCC AAA GGA AGT ATT TTT GAG GCT CTT CCA AAT GAC	1661
156 S D A M H Y L L S K G S I F E A L P N D	175
1662 AAT TAC CTT CAG ATT TCT GGC ATA CCA CTT TTT AAA AAT AAT GTG TTT GAG GAA ACT GTG	1721
176 N Y L Q I S G I P L F K N N V F E E T V	195
1722 TCA AAA AAA AGA AAG CGA ACC ATT GAA ACA TCC ATT ACT CAA AAT AAA AGC GCC CGC AAA	1781
196 S K K R K R T I E T S I T Q N K S A R K	215
1782 GAA GTT TCC TGG AAT AGC ATT TCA ATT AGT AGG TTT AGC ATT TTT TAC AGG TCA TCC TAT	1841
216 E V S W N S I S R F S I F Y R S S Y	235
1842 AAG AAG TTT AAG CAA G gtaactaatactgttatccatataactatTTT AT CTA TAT TTT AAC	
1907	
236 K K F K Q D	L Y F N 245
1908 TTA CAC TCT ATT TGT GAT CGG AAC ACA GTA CAC ATG TGG CTT CAA TGG ATT TTT CCA AGG	1967
246 L H S I C D R N T V H M W L Q W I F P R	265
1968 CAA TTT GGA CTT ATA AAC GCA TTT CAA GTG AAG CAA TTG CAC AAA GTG ATT CCA CTG GTA	2027
266 Q F G L I N A F Q V K Q L H K V I P L V	285
2028 TCA CAG AGT ACA GTT GTG CCC AAA CGT CTC CTA AAG GTA TAC CCT TTA ATT GAA CAA ACA	2087
286 S Q S T V V P K R L L K V Y P L I E Q T	305
2088 GCA AAG CGA CTC CAT CGT ATT TCT CTA TCA AAA GTT TAC AAC CAT TAT TGC CCA TAT ATT	2147
306 A K R L H R I S L S K V Y N H Y C P Y I	325
2148 GAC ACC CAC GAT GAT GAA AAA ATC CTT AGT TAT TCC TTA AAG CCG AAC CAG GTG TTT GCG	2207
326 D T H D D E K I L S Y S L K P N Q V F A	345
2208 TTT CTT CGA TCC ATT CTT GTT CGA GTG TTT CCT AAA TTA ATC TGG GGT AAC CAA AGG ATA	2267
346 F L R S I L V R V F P K L I W G N Q R I	365
2268 TTT GAG ATA ATA TTA AAA G gtattgtataaaaatttattaccactaacgatttaccag AC CTC GAA ACT	2336
366 F E I I L K D	L E T 375

FIGURE 46 (cont.)

2337	TTC TTG AAA TTA TCG AGA TAC GAG TCT TTT AGT TTA CAT TAT TTA ATG AGT AAC ATA AAG	2396
376	F L K L S R Y E S F S L H Y L M S N I K	395
2397	gtatatatgccaaatttttaccattaattaacaatcg ATT TCA GAA ATT GAA TGG CTA GTC CTT GGA	2465
396	I S E I E W L V L G	405
2466	AAA AGG TCA AAT GCG AAA ATG TGC TTA AGT GAT TTT GAG AAA CGC AAG CAA ATA TTT GCG	2525
406	K R S N A K M C L S D F E K R K Q I F A	425
2526	GAA TTC ATC TAC TGG CTA TAC AAT TCG TTT ATA ATA CCT ATT TTA CAA TCT TTT TTT TAT	2585
426	E F I Y W L Y N S F I I P I L Q S F F Y	445
2586	ATC ACT GAA TCA AGT GAT TTA CGA AAT CGA ACT GTT TAT TTT AGA AAA GAT ATT TGG AAA	2645
446	I T E S S D L R N R T V Y F R K D I W K	465
2646	CTC TTG TGC CGA CCC TTT ATT ACA TCA ATG AAA ATG GAA GCG TTT GAA AAA -ATA AAC GAG	2705
466	L L C R P F I T S M K M E A F E K I N E	485
2706	gtatTTaaagtatTTTgcAAAAagctaataTTTcag AAC AAT GTT AGG ATG GAT ACT CAG AAA ACT	2775
486	N N V R M D T Q K T	495
2776	ACT TTG CCT CCA GCA GTT ATT CGT CTA TTA CCT AAG AAG AAT ACC TTT CGT CTC ATT ACG	2835
496	T L P P A V I R L L P K K N T F R L I T	515
2836	AAT TTA AGA AAA AGA TTC TTA ATA AAG gtattaattttggcatcaatgtacttacttctaattat	2906
516	N L R K R F L I K	524
2907	ttagcag ATG GGT TCA AAC AAA AAA ATG TTA GTC AGT ACG AAC CAA ACT TTA CGA CCT GTG	2967
525	M G S N K K M L V S T N Q T L R P V	542
2968	GCA TCG ATA CTG AAA CAT TTA ATC AAT GAA GAA AGT AGT GGT ATT CCA TTT AAC TTG GAG	3027
543	A S I L K H L I N E E S S G I P F N L E	562
3028	GTT TAC ATG AAG CTT CTT ACT TTT AAG AAG GAT CTT CTT AAG CAC CGA ATG TTT GG gtaat	3088
563	V Y M K L L T F K K D L L K H R M F G	581
3089	tatataatgcgcgattcctcattattaattttcag G CGT AAG AAG TAT TTT GTA CGG ATA GAT ATA	3155
582	R K K Y F V R I D I	591
3156	AAA TCC TGT TAT GAT CGA ATA AAG CAA GAT TTG ATG TTT CGG ATT GTT AAA AAG AAA CTC	3215
592	K S C Y D R I K Q D L M F R I V K K K L	611
3216	AAG GAT CCC GAA TTT GTA ATT CGA AAG TAT GCA ACC ATA CAT GCA ACA AGT GAC CGA GCT	3275
612	K D P E F V I R K Y A T I H A T S D R A	631
3276	ACA AAA AAC TTT GTT AGT GAG GCG TTT TCC TAT T gtaagttatTTTcattgaaatttttaacaa	3343
632	T K N F V S E A F S Y F	643
3344	attcttttttag TT GAT ATG GTG CCT TTT GAA AAA GTC GTG CAG TTA CTT TCT ATG AAA ACA	3405
644	D M V P F E K V V Q L L S M K T	659
3406	TCA GAT ACT TTG TTT GTT GAT TTT GTG GAT TAT TGG ACC AAA AGT TCT TCT GAA ATT TTT	3465
660	S D T L F V D F V D Y W T K S S S E I F	679
3466	AAA ATG CTC AAG GAA CAT CTC TCT GGA CAC ATT GTT AAG gtataccaattgttgaattgtataaca	3532
680	K M L K E H L S G H I V K	692

FIGURE 46 (cont.)

3533 ctaatgaaaactag ATA GGA AAT TCT CAA TAC CTT CAA AAA GTT GGT ATC CCT CAG GGC TCA	3593
693 : I G N S Q Y L Q K V G I P Q G S	708
3594 ATT CTG TCA TCT TTT TTG TGT CAT TTC TAT ATG GAA GAT TTG ATT GAT GAA TAC CTA TCG	3653
709 I L S S F L C H F Y M E D L I D E Y L S	728
3654 TTT ACG AAA AAG AAA GGA TCA GTG TTG TTA CGA GTA GTC GAC GAT TTC CTC TTT ATA ACA	3713
729 F T K K K G S V L L R V V D D F L F I T	748
3714 GTT AAT AAA AAG GAT GCA AAA AAA TTT TTG AAT TTA TCT TTA AGA G gtgagttgtgtcattcc	3777
749 V N K K D A K K F L N L S L R G	764
3778 taagttctaaccgttgaag GA TTT GAG AAA CAC AAT TTT TCT ACG AGC CTG GAG AAA ACA GTA	3840
765 F E K H N F S T S L E K T V	778
3841 ATA AAC TTT GAA AAT AGT AAT GGG ATA ATA AAC AAT ACT TTT TTT AAT GAA AGC AAG AAA	3900
779 I N F E N S N G I I N N T F F N E S K K	798
3901 AGA ATG CCA TTC TTC GGT TCT GTG AAC ATG AGG TCT CTT GAT ACA TTG TTA GCA TGT	3960
7799 R M P F F G F S V N M R S L D T L L A C	818
3961 CCT AAA ATT GAT GAA GCC TTA TTT AAC TCT ACA TCT GTA GAG CTG ACG AAA CAT ATG GGG	4020
7819 P K I D E A L F N S T S V E L T K H M G	838
4021 AAA TCT TTT TTT TAC AAA ATT CTA AG gtatactgtgtactgaaataatagctgacaaataatcag A TCG	4089
7839 K S F F Y K I L R S	848
4090 AGC CTT GCA TCC TTT GCA CAA GTA TTT ATT GAC ATT ACC CAC AAT TCA AAA TTC AAT TCT	4149
7849 S L A S F A Q V F I D I T H N S K F N S	868
4150 TGC TGC AAT ATA TAT AGG CTA GGA TAC TCT ATG TGT ATG AGA GCA CAA GCA TAC TTA AAA	4209
7869 C C N I Y R L G Y S M C M R A Q A Y L K	888
4210 AGG ATG AAG GAT ATA TTT ATT CCC CAA AGA ATG TTC ATA ACG G gtgagttacttattttactaga	4274
7889 R M K D I F I P Q R M F I T D	903
4275 aaagtcatataattaaaccttag AT CTT TTG AAT GTT ATT GGA AGA AAA ATT TGG AAA AAG TTG GCC	4339
904 L L N V I G R K I W K K L A	917
4340 GAA ATA TTA GGA TAT ACG AGT AGG CGT TTC TTG TCC TCT GCA GAA GTC AAA TG gtacgtgtc	4401
918 E I L G Y T S R R F L S S A E V K W	935
4402 ggctcgagacttcagcaatattgacacatcag G CTT TTT TGT CTT GGA ATG AGA GAT GGT TTG AAA	4468
936 L F C L G M R D G L K	946
4469 CCC TCT TTC AAA TAT CAT CCA TGC TTC GAA CAG CTA ATA TAC CAA TTT CAG TCA TTG ACT	4528
947 P S F K Y H P C F E Q L I Y Q F Q S L T	966
4529 GAT CTT ATC AAG CCG CTA AGA CCA GTT TTG CGA CAG GTG TTA TTT TTA CAT AGA AGA ATA	4588
967 D L I K P L R P V L R Q V L F L H R R I	986
4589 GCT GAT TAA tgtcatttcaatttattatatacatccttattactgggtcttaacaatattactaagtata	4665
987 A D *	989

FIGURE 46 (cont.)

4666 gctgaccccaaagcaagcatactataggattctagtaaagtaaaattaatctcgttat tagtttatttgacttgtct 4745
4746 ttatccttatactttaaagaagattgacagtgggtgctgactactgcccacatgcccattaaacgggagtggtaaaca 4825
4826 ttaaaagtaatacatgaggctaattcccttcatttagaataaggaaagtggttctataatgaataatgcccacta 4905
4906 atgaaaaagacgaagattatctctaaacaaggggattaagcatatccgaaggaaaagagagtaatataccagtgtt 4985
4986 gttgaagaaagcaaggataatttggacaagcttctgcagatgacaggctaaatttggtaaccgaatttggtaaaagc 5065
5066 cccaggttatccatggccggccttgcattactgagacgaaaactaaggatagttgaataactaatacgctattta 5145
5146 atgtcttatataagggtttgttttgcattcaatttgcattggtaaaaagaaatagtgttaagccattattggat 5225
5226 tcggaaatagccaaatttcttggctctcaaagcggaaagtctaaagaacttattgaagcttatgaggcttcaaaaactcc 5305
5306 tcctgatttaaggaggaatcttccaccgatgagggaaatggatagcttatcagctgctgaggagaagcctaattttgc 5385
5386 aaaaaagaaaatcattggagacatctttgatgaatcagatgcggagagtatctccagcggatcctgatgtcaata 5465
5466 acttctatttctgaaatgtatggtcctactgtcgcttcgacttctcgtagctctacgcagttaaagtgaccaaaggtaacc 5544

FIGURE 47

1
met ser val tyr val val glu leu leu
GCCAAGTTCTGCCTGGCTG ATG AGT GTG TAC GTC GTC GAG CTG CTC

10 20
arg ser phe phe tyr val thr glu thr phe gln lys asn arg
AGG TCT TTC TTT TAT GTC ACG GAG ACC ACG TTT CAA AAG AAC AGG

30
leu phe phe tyr arg lys ser val trp ser lys leu gln ser ile
CTC TTT TTC TAC CGG AAG AGT GTC TGG AGC AAG TTG CAA AGC ATT

40 50
gly ile arg gln his leu lys arg val gln leu arg glu leu ser
GGA ATC AGA CAG CAC TTG AAG AGG GTG CAG CTG CGG GAG CTG TCG

60
glu ala glu val arg gln his arg glu ala arg pro ala leu leu
GAA GCA GAG GTC AGG CAG CAT CGG GAA GCC AGG CCC GCC CTG CTG

70 80
thr ser arg leu arg phe ile pro lys pro asp gly leu arg pro
ACG TCC AGA CTC CGC TTC ATC CCC AAG CCT GAC GGG CTG CGG CCG

90
ile val asn met asp tyr val val gly ala arg thr phe arg arg
ATT GTG AAC ATG GAC TAC GTC GTG GGA GCC AGA ACG TTC CGC AGA

100 110
glu lys ala glu arg leu thr ser arg val lys ala leu phe
GAA AAG ARG GCC GAG CGT CTC ACC TCG AGG GTG AAG GCA CTG TTC

120
ser val leu asn tyr glu arg ala arg arg pro gly leu leu gly
AGC GTG CTC AAC TAC GAG CGG CGG CGC CCC GGC CTC CTG GGC

130 140
ala ser val leu gly leu asp asp ile his arg ala trp arg thr
GCC TCT GTG CTG GGC CTG GAC GAT ATC CAC AGG GCC TGG CGC ACC

150
phe val leu arg val arg ala gln asp pro pro pro glu leu tyr
TTC GTG CTG CGT GTG CGG GCC CAG GAC CCG CCT GAG CTG TAC

160 170
phe val lys val asp val thr gly ala tyr asp thr ile pro gln
TTT GTC AAG GTG GAT GTG ACG GGC GCG TAC GAC ACC ATC CCC CAG

180
asp arg leu thr glu val ile ala ser ile ile lys pro gln asn
GAC AGG CTC ACG GAG GTC ATC GCC AGC ATC AAA CCC CAG AAC

FIGURE 47 (cont.)

190 200
thr tyr cys val arg arg tyr ala val val gln lys ala ala met
ACG TAC TGC GTG CGT CGG TAT GCC GTG GTC CAG AAG GCC GCC ATG

210
gly thr ser ala arg pro ser arg ala thr ser tyr val gln cys
GGC ACG TCC GCA AGG CCT TCA AGA GCC ACG TCC TAC GTC CAG TGC

220 230
gln gly ile pro gln gly ser ile leu ser thr leu leu cys ser
CAG GGG ATC CCG CAG GGC TCC ATC CTC TCC ACG CTG CTC TGC AGC

240
leu cys tyr gly asp met glu asn lys leu phe ala gly ile arg
CTG TGC TAC GGC GAC ATG GAG AAC AAG CTG TTT GCG GGG ATT CGG

250 260
arg asp gly leu leu leu arg leu val asp asp phe leu leu val
CGG GAC GGG CTG CTC CTG CGT TTG GTG GAT GAT TTC TTG TTG GTG

270
thr pro his leu thr his ala lys thr phe leu arg thr leu val
ACA CCT CAC CTC ACC CAC GCG AAA ACC TTC CTC AGG ACC CTG GTC

280 290
arg gly val pro glu tyr gly cys val val asn leu arg lys thr
CGA GGT GTC CCT GAG TAT GGC TGC GTG AAC TTG CGG AAG ACA

300
val val asn phe pro val glu asp glu ala leu gly gly thr ala
GTG GTG AAC TTC CCT GTA GAA GAC GAG GCC CTG GGT GGC ACG GCT

310 320
phe val gln met pro ala his gly leu phe pro trp cys gly leu
TTT GTT CAG ATG CCG GCC CAC GGC CTA TTC CCC TGG TGC GGC CTG

330
leu leu asp thr arg thr leu glu val gln ser asp tyr ser ser
CTG CTG GAT ACC CGG ACC CTG GAG GTG CAG AGC GAC TAC TCC AGC

340 350
tyr ala arg thr ser ile arg ala ser leu thr phe asn arg gly
TAT GCC CGG ACC TCC ATC AGA GCC AGT CTC ACC TTC AAC CGC GGC

360
phe lys ala gly arg asn met arg arg lys leu phe gly val leu
TTC AAG GCT GGG AGG AAC ATG CGT CGC AAA CTC TTT GGG GTC TTG

370 380
arg leu lys cys his ser leu phe leu asp leu gln val asn ser
CGG CTG AAG TGT CAC AGC CTG TTT CTG GAT TTG CAG GTG AAC AGC

FIGURE 47 (cont.)

390

leu gln thr val cys thr asn ile tyr lys ile leu leu leu gln
CTC CAG ACG GTG TGC ACC AAC ATC TAC AAG ATC CTC CTG CTG CAG

400

410

ala tyr arg phe his ala cys val leu gln leu pro phe his gln
GCG TAC AGG TTT CAC GCA TGT GTG CTG CAG CTC CCA TTT CAT CAG

420

gln val trp lys asn pro his phe ser cys ala ser ser leu thr
CAA GTT TGG AAG AAC CCA CAT TTT TCC TGC GCG TCA TCT CTG ACA

430

440

arg leu pro leu leu leu his pro glu ser gln glu arg arg asp
CGG CTC CCT CTG CTA CTC CAT CCT GAA AGC CAA GAA CGC AGG GAT

450

val ala gly gly gln gly arg arg pro ser ala leu arg gly
GTC GCT GGG GGC CAA GGG CGC CGC CGG CCC TCT GCC CTC CGA GGC

460

470

arg ala val ala val pro pro ser ile pro ala gln ala asp ser
CGT GCA GTG GCT GTG CCA CCA AGC ATT CCT GCT CAA GCT GAC TCG

480

thr pro cys his leu arg ala thr pro gly val thr gln asp ser
ACA CCG TGT CAC CTA CGT GCC ACT CCT GGG GTC ACT CAG GAC AGC

490

500

pro asp ala ala glu ser glu ala pro gly asp asp ala asp cys
CCA GAC GCA GCT GAG TCG GAA GCT CCC GGG GAC GAC GCT GAC TGC

510

pro gly gly arg ser gln pro gly thr ala leu arg leu gln asp
CCT GGA GGC CGC AGC CAA CCC GGC ACT GCC CTC AGA CTT CAA GAC

520

530

his pro gly leu met ala thr arg pro gln pro gly arg glu gln
CAT CCT GGA CTG ATG GCC ACC CGC CCA CAG CCA GGC CGA GAG CAG

540

thr pro ala ala leu ser arg arg ala tyr thr ser gln gly gly
ACA CCA GCA GCC CTG TCA CGC CGG GCT TAT ACG TCC CAG GGA GGG

550

560

arg gly gly pro his pro gly leu his arg trp glu ser glu ala
AGG GGC GGC CCA CAC CCA GGC CTG CAC CGC TGG GAG TCT GAG GCC

564

OP

TGA GTGAGTGTTGGCCGAGGCCTGCATGTCCGGCTGAAGGCTGAGTGTCCGGCTGAGGC
CTGAGCGAGTGTCAGCCAAGGGCTGAGTGTCCAGCACACCTGCGTTTCACTTCCCCAC

FIGURE 47 (cont.)

FIGURE 48

Motif -1
 Ep p123 ...LVVSLIRCFYVTEQQKSYSKT...
 Sp Tez1 ...FIIPILQSFFYITESSDLRNRT...
 Sc Est2 ...LIPKIIQTFFYCTEISSTVTIV...
 Hs TCP1 ...YVVELLRSFFYVTETTFQKNRL...
 consensus FFY TE

K
 Motif 0 p hhh K hR h R
 Ep p123 ...KSLGFAPGKLRLIPKKT--TFRPIMTFNKKIV...
 Sp Tez1 ...QKTTLPPAVIRLLPKKN--TFRLITNLKRFL...
 Sc Est2 ...TLSNFNHSKMRIPKKSNNEFRIIAIPCRGAD...
 Hs TCP1 ...ARPALLTSRLRFIPKPD--GLRPIVNMDYVVG...
 consensus R PK RI

AF
 Motif A h hDh GY h
 Ep p123 ...PKLFATMDIEKCYDSVNREKLSTFLK...
 Sp Tez1 ...RKKYFVRIDIKSCYDRIKQDLMFRJVK...
 Sc Est2 ...PELYFMKFDVKSCYDSIPRMECMRILK...
 Hs TCP1 ...PELYFVKVDVTGAYDTIPQDRLTEVIA...//...
 consensus F D YD

hPQG pS hh
 Motif B Ep p123 ...NGKFYKQTKGIPQGLCVSSILSSFYA...
 Sp Tez1 ...GNSQYLQKVGVIPQGSILSSFLCHFYME...
 Sc Est2 ...EDKCYIREDGLFQGSSLSAPIVDLVYD...
 Hs TCP1 ...RATSYVQCQGIPQGSILSTLLCSLCYG...
 consensus G QG S

Y
 Motif C h F DDhhh
 Ep p123 ...PNVNLLMRLTDYLLITTQENN...
 Sp Tez1 ...KKGSVLLRVVDDFLFITVNKKD...
 Sc Est2 ...SQDTLILKLAADDFLIISTDQQQ...
 Hs TCP1 ...RRDGLLLRLVDDFLLVTPHLTH...
 consensus DD L

Gh h cK
 Motif D Ep p123 ...NVSRENGFKFNMKKL...
 Sp Tez1 ...LNLSLRGFEKHNFS...
 Sc Est2 ...KKLAMGGFQKYNAKA...
 Hs TCP1 ...LRTLVRGVPEYGCVV...
 consensus G

FIGURE 49

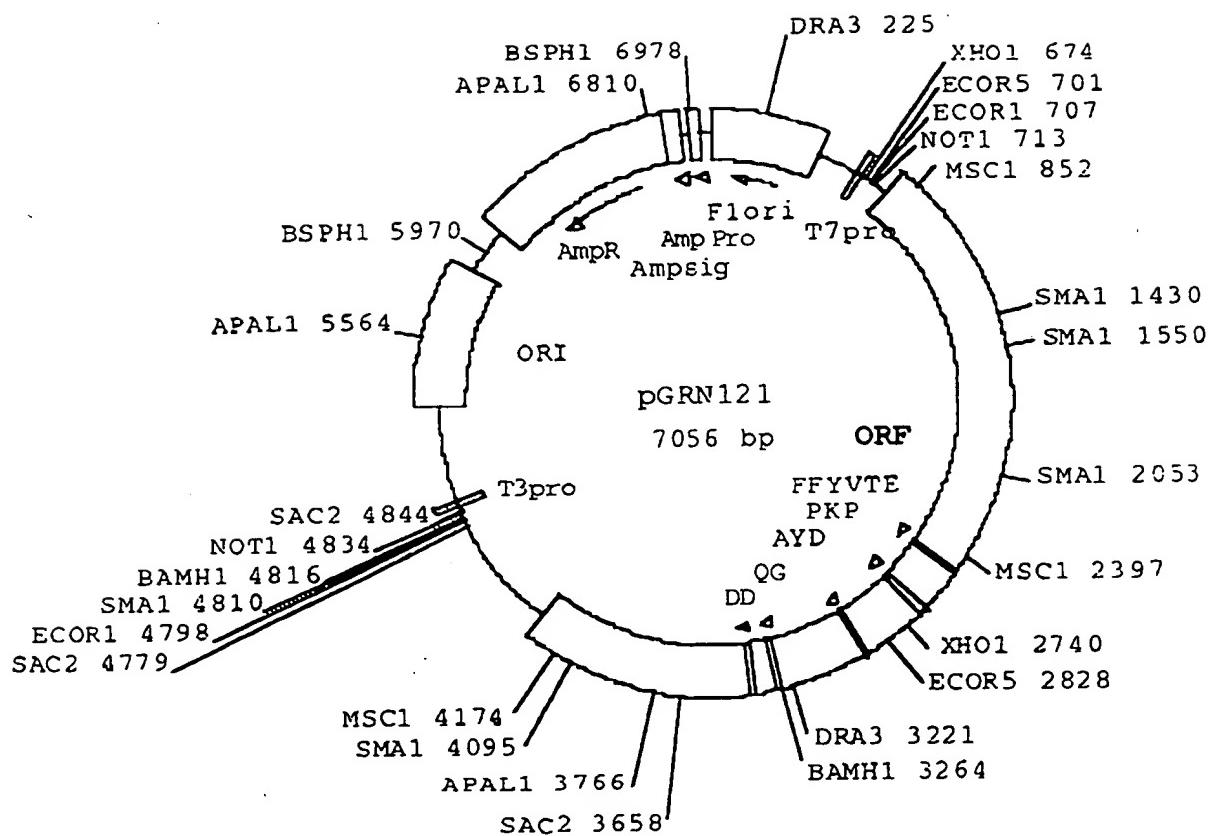


FIGURE 50

1 GCAGCGCTGC GTCCTGCTGC GCACGTGGGA AGCCCTGGCC CCGGCCACCC
51 CCGCGATGCC GCGCGCTCCC CGCTGCCGAG CGTGCGCTC CCTGCTGCGC
101 AGCCACTACC GCGAGGTGCT GCCGCTGGCC ACGTCGTGC GGCGCTGGG
151 GCCCCAGGGC TGGCGGCTGG TGCAAGCGGG GGACCCGGCG GCTTCCGCG
201 CGNTGGTGGC CCANTGCNTG GTGTGCGTGC CCTGGGANGN ANGGCNGCCC
251 CCCGCCGCC CCTCCTTCCG CCAGGTGTCC TGCCTGAANG ANCTGGTGGC
301 CCGAGTGTG CANANGCTGT GCGANCGGG CGCGAANAAC GTGCTGGCCT
351 TCGGCTTCGC GCTGCTGGAC GGGGCCCGCG GGGGCCCCCG CGAGGCCTTC
401 ACCACCAGCG TGCGCAGCTA CCTGCCAAC ACGGTGACCG ACGCACTGCG
451 GGGGAGCGGG GCGTGGGGGC TGCTGCTGCG CCGCGTGGGC GACGACGTGC
501 TGGTTCACCT GCTGGCACGC TGCGCGNTNT TTGTGCTGGT GGNTCCCAGC
551 TGCGCCTACC ANGTGTGCGG GCCGCCGCTG TACCAGCTCG GCGCTGCNAC
601 TCAGGCCCGG CCCCCGCCAC ACGCTANTGG ACCCGAANGC GTCTGGGATC
651 CAACGGGCCT GGAACCATAG CGTCAGGGAG GCCGGGGTCC CCCTGGGCTG
701 CCAGCCCCGG GTGCGAGGAG GCGCGGGGGC AGTGCCAGCC GAAGTCTGCC
751 GTTGCCCAAG AGGCCCAGGC GTGGCGCTGC CCCTGAGCCG GAGCGGACGC
801 CCGTTGGGCA GGGGTCTGG GCCCACCGG GCAGGACGCC TGGACCGAGT
851 GACCGTGGTT TCTGTGTGGT GTCACCTGCC AGACCCGCCG AAGAAGGCCAC
901 CTCTTGAGG GGTGCGCTCT CTGGCACGCC CCACTCCCAC CCATCCGTGG
951 GCCGCCAGCA CCACGCCGGC CCCCCATCCA CATGCCGCC ACCACGTCC
1001 GGGACACGCC TTGTCCCCCG GTGTACGCCAG AGACCAAGCA CTTCCCTTAC
1051 TCCTCAGGCCG ACAAGNACAC TGCGNCCCTC CTTCTACTC AATATACTG
1101 AGGCCCAGCC TGACTGGCGT TCAGGGAGGTT CGTGGAGACA NTCTTCTGG
1151 TTCCAGGCCCT TGGATGCCAG GATTCCCCGC AGGTTGCCCC GCCTGCCCA
1201 GCGNTACTGG CAAATGCCGC CCCTGTTCT GGAGCTGCTT GGGAACACAG
1251 CGCAGTGCCCT ACAGGGGTG TTCCTCAAGA CGCACTGCC GCTGCGAGCT
1301 GCGGTACCCC CAGCAGCCGG TGTCTGTGCC CGGGAGAACG CCCAGGGCTC
1351 TGTGGCGGCC CCCGAGGAGG AGGAACACAG ACCCCCCGTG CCTGGTGCAG
1401 CTGCTCCGCC AGCACACGCCAG CCCCTGGCAG GTGTACGGCT TCGTGCAGGGC
1451 CTGCCCTGCC CGGCTGGTGC CCCCCAGGCCCT CTGGGGCTCC AGGCACAACG
1501 AACGCCGCTT CCTCAGGAAC ACCAAGAAGT TCATCTCCCT GGGGAAGCAT
1551 GCCAAGCTCT CGCTGCAGGA GCTGACGTGG AAGATGAGCG TGCAGGGACTG
1601 CGCTTGGCTG CGCAGGAGCC CAGGGGTTGG CTGTGTTCCG GCCGCAGAGC
1651 ACCGTCTGCCG TGAGGAGATC CTGGCCAAGT TCCTGCACTG GCTGATGAGT
1701 GTGTACGTGCG TCGAGCTGCT CAGGTCTTTC TTTTATGTCA CGGAGACCCAC
1751 GTTTAAAAG AACAGGCTCT TTTTCTACCG GAAGAGTGTGTC TGGAGCAAGT
1801 TGCAAAGCAT TGAATCAGA CAGCACTTGA AGAGGGTGCA GCTGCGGGAG
1851 CTGTCGGAAG CAGAGGTCAAG GCAGCATCGG GAAGCCAGGC CGGCCCTGCT
1901 GACGTCCAGA CTCCGCTTCA TCCCCAAGCC TGACGGGCTG CGGCCGATTG
1951 TGAACATGGA CTACGTCGTG GGAGCCAGAA CGTTCCGCAG AGAAAAGAGG
2001 GCCGAGCGTC TCACCTCGAG GGTGAAGGCA CTGTTCAGCG TGCTCAACTA
2051 CGAGCGGGCG CGGCGCCCCCG GCCTCCTGGG CGCCTCTGTG CTGGGCCTGG
2101 ACGATAATCCA CAGGGCCTGG CGCACCTTCA TGCTGCGTGT CGGGGCCAG
2151 GACCCGCCGC CTGAGCTGTA CTTTGTCAAG GTGGATGTGA CGGGCGCGTA
2201 CGACACCACCC CCCCCAGGACA GGCTCACGGA GGTCATCGCC AGCATCATCA
2251 AACCCCCAGAA CACGTACTGC GTGGCGTGGT ATGCCGTGGT CCAGAAGGCC

FIGURE 50 (cont.)

2301 GCCCATGGGC ACGTCCGCAA GGCTTCAAG AGCCACGTCT CTACCTTGAC
2351 AGACCTCCAG CCGTACATGC GACAGTTGT GGCTCACCTG CAGGANAACA
2401 GCCCGCTGAG GGATGCCGTC GTCATCGAGC AGAGCTCCTC CCTGAATGAG
2451 GCCAGCAGTG GCCTCTCGA CGTCTTCCTA CGCTTCATGT GCCACCACGC
2501 CGTGCACATC AGGGGCAAGT CCTACGTCCA GTGCCAGGGG ATCCCGCAGG
2551 GCTCCATCCT CTCCACGCTG CTCTGCAGCC TGTGCTACGG CGACATGGAG
2601 AACAAAGCTGT TTGCGGGGAT TCGCGGGAC GGGCTGCTCC TGCCTTGGT
2651 GGATGATTTT TTGTTGGTGA CACCTCACCT CACCCACGCG AAAACCTTCC
2701 TCAGGACCCCT GGTCGAGGT GTCCCTGAGT ATGGCTGCGT GGTGAACTTG
2751 CGGAAGACAG TGGTGAACCT CCCTGTAGAA GACGAGGCCG TGGGTGGCAC
2801 GGCTTTGTT CAGATGCCGG CCCACGGCCT ATTCCCCTGG TGCGGCCTGC
2851 TGCTGGATAAC CCGGACCCCTG GAGGTGCAGA GCGACTACTC CAGCTATGCC
2901 CGGACCTCCA TCAGAGCCAG TCTCACCTTC AACCGCGGCT TCAAGGCTGG
2951 GAGGAACATG CGTCGCAAAC TCTTGGGGT CTTGCGGCTG AAGTGTACA
3001 GCCTGTTCT GGATTGCAAG GTAACACAGCC TCCAGACGGT GTGCACCAAC
3051 ATCTACAAGA TCCTCCTGCT GCAGGCGTAC AGGTTTCACG CATGTGTGCT
3101 GCAGCTCCA TTTCATCAGC AAGTTGGAA GAACCCCACA TTTTCCCTGC
3151 GCGTCATCTC TGACACGGCC TCCCTCTGCT ACTCCATCCT GAAAGCCAAG
3201 AACGCAGGGTA TGTCGCTGGG GGCAAGGGC GCCGCCGGCC CTCTGCCCTC
3251 CGAGGCCGTG CAGTGGCTGT GCCACCAAGC ATTCTGCTC AAGCTGACTC
3301 GACACCGTGT CACCTACGTG CCACCTCTGG GGTCACTCAG GACAGCCCAG
3351 ACGCAGCTGA GTCGGAAGCT CCCGGGGACG ACGCTGACTG CCCTGGAGGC
3401 CGCAGCCAAC CCGGCACTGC CCTCAGACTT CAAGACCATC CTGGACTGAT
3451 GCCCACCCGC CCACAGCCAG GCCGAGAGCA GACACCAGCA GCCCTGTCAC
3501 GCCGGGCTCT ACGTCCCAGG GAGGGAGGGG CGGCCACAC CCAGGCCCGC
3551 ACCGCTGGGA GTCTGAGGCC TGAGTGAGTG TTTGGCCGAG GCCTGCATGT
3601 CCGGCTGAAG GCTGAGTGTC CGGCTGAGGC CTGAGCGAGT GTCCAGCCAA
3651 GGGCTGAGTG TCCAGCACAC CTGCCGTCTT CACTCCCCA CAGGCTGGCG
3701 CTCGGCTCCA CCCCCAGGGCC AGCTTTCCCT CACCAGGAGC CCGGCTTCCA
3751 CTCCCCACAT AGGAATAGTC CATCCCCAGA TTCGCCATTG TTCACCCCTC
3801 GCCCTGCCCT CCTTGCCTT CCACCCCCAC CATCCAGGTG GAGACCTGA
3851 GAAGGACCCCT GGGAGCTCTG GGAATTGGA GTGACCAAAG GTGTGCCCTG
3901 TACACAGGCG AGGACCCCTGC ACCTGGATGG GGGTCCCTGT GGGTCAAATT
3951 GGGGGGAGGT GCTGTGGAG TAAAATACTG AATATATGAG TTTTCAGTT
4001 TTGAAAAAAA AAAAAAAAAA AAAAAAAAAA

FIGURE 51

FIGURE 51 (cont.)

CCTGCCCAACAGGGTGACCGAAGCCTGCGGGGAGGGGGTGGGGCTGTCTCTGCG
 421 +-----+-----+-----+-----+-----+-----+-----+-----+-----+
 CCACGGGTTGACCACTGGCTGCGTGACGCCCTTCGCCCGAACCGAGGACGAAAGC
 P A Q H G D R R T A G E F G V G A A A A
 L P N T V T D A I R G S G A W G L L R -
 C P T R * F T H C C C A G R G G G C C A -
 CGGGTGGGGAGGAGCTGGTGGTTCACCIAGCTGGCACCCAGGCGGCTGCTGCTGCT
 481 +-----+-----+-----+-----+-----+-----+-----+-----+-----+
 CCCACCCGGCTGCTGACGACCAAGTGGAGGACCGTGGGAAGGCGCNANAACACGACCA
 P R C R R R A G S P A G T L R ? ? C A G -
 R V G D D V L V H L L A R C A ? F V L V -
 A W A T T C W F T C W H A A R ? L C W W -
 GGNCCCCACCTGGCGGCTTACCAAGTGTGGGGGGCGCGCTGTACCAAGCTGGGGCTGCGAC
 541 +-----+-----+-----+-----+-----+-----+-----+-----+-----+
 OCNAGGGTGGACGGGATGGTNCACACCCCCGGCGGGGACATGGTACACCCCCGGACCGT
 G S Q L R L P ? V R A A A V P A R R C ? -
 ? P S C A Y ? V C G P P L Y Q L G A A T -
 ? F A A P T ? C A G R R C T S S A L Y L -
 TCAGGCCCCCCCCGGCACACCCCTANTGGACCCCCAAGGCTCTGGATCAAACGGGCT
 601 +-----+-----+-----+-----+-----+-----+-----+-----+-----+
 ACTGGGGCGGGGGGGGGATGGATACCTGGCTTNGCAGACCCCTAGGTTGGGAGA
 S C P A P A T R ? W T R ? R L G S N G P -
 Q A R P P F H A ? G P E ? V W D P T G L -
 R P G P R H T L ? D F ? A S G I Q R A W -
 GGAACCATAGGGTAGGGAGGGCGGCGGCTGGGGCTGGGGCTCCACCCACCCGGGTOCCACCG
 661 +-----+-----+-----+-----+-----+-----+-----+-----+-----+
 CCTGGTATGGAGGCGGCTCGGGGGGGGAGGGGAGGGGACGGGAGGGCTGGGGCCACGGCTGG
 G T I A S C R P C S P W A A S P G C E E -
 E P * R Q G G R C P P G L P A F G A R R -
 N H S V R E A G V P L G C Q P R V R G G -
 GGCACGCCCCAGTGGAGGCGGAGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGG
 721 +-----+-----+-----+-----+-----+-----+-----+-----+-----+
 CGGGGGGGGGTGGAGGTCGGTGGCTGGGGCTGGGGGGGGGGGGGGGGGGGGGGGGGGGG
 A R G U C Q P K S A V A V A Q E A Q A W R C
 R G G S A S R S L P L P K R P R R G A A -
 A G A V F A E V C R C F R G P G V A L P -
 CCCAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG
 781 +-----+-----+-----+-----+-----+-----+-----+-----+-----+
 GGGACTGGGGCTGGGGCTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG
 P - A C A D A R W A C V L G F F G Q D A -
 P E P E R T P V G Q G S W A H P G R T F -
 L S R S G R P L G R G P G F T R A C R L -

FIGURE 51 (cont.)

65
60
55
50
45
40
35
30
25
20
15
10
5
0

841 TGGACCGAGTCACCGTGGTTCTGTGCGGTCACCTCCACACCCUCUCAAGAACCCAC
 900
 ACCTGGCTACTGGCACCAAGACACACACAGTGGACGGTCGGGGCGCTTCGGCG
 a W T E * P W F L C G V T C Q T R R R S H -
 b G P S D R G F C V V S P A R P A E E A T
 c D R V T V V S V W C H L P D P P K K P F -
 901 CTCTTGGAGGCTGGCTCTCTGGCAAGGOGOCACCTCCACACCTGGGGGGCGCGA
 960
 GAGAAACCTCCACCGGAGAGACCGTGCGCGGTGAGGGTGGCTAGGCACCCGGCGTCGT
 a L F G G C A L W H A P L F P I R G F E A
 b S L E G A L S C T R H S H P S V G R Q H -
 c L W R V R S L A R A T P T H P W A A S T -
 961 CCAO3GGGGCCCCCATCCACATCGGGCCACCAAGTCCAGGACACGCCCTGGCCCCG
 1020
 CCTGGCCCCGGGGCTAGGTGTAGGGCGGGTGGTGCAGGAACCTGTGGGGAAACGGGGC
 a P R C P P I H I A A T T S W D T F C P P -
 b H A G P P S T S R P P R P G T R L V P E
 c T R A F H P H R G H H V L G H A L S P G -
 1021 GTGTAAACGGAGACCAAGCACTTCCTACTCTCAGGCAGACAGNACACTGGNNCCCT
 1080
 CACATGGCTCTGGTGTGAAGGAGATGAGGGAGTCCCTGTCTGAGCGCNGGCAG
 a V Y A E T K H F L Y S S C D K ? T A ? L -
 b C T P R P S T S S T P Q A T ? T L R P S -
 c V R E D Q A L P L L L R R Q ? H C ? P P -
 1081 CTTOCTACTCAATATACTGAGGCCAGGCTGACTGGCGTTGGAGGTTCGTGGAGAC
 1140
 GAAGGATGAGTTATATAGACCTGGGGTGGACTGACCGCAAGCCCTCCAAAGCACCTCT
 a L P T Q Y I * G P A * L A F G R F V E T -
 b F L L N I S E A Q P D W R S G G S W R ? -
 c S Y S I Y L R P S L T G V R E V R G D ? -
 1141 NTCTTCTGGTTCAGGCCCTGGATGCCAGGATTCCCGCAGGTGGCGCCGGCTGGCGA
 1200
 NAGAAAGACCAAGGAGGGAAACCTGGGCTCCAGGCGTCCAAACGGGGCGAACGGGGT
 a ? P L V P G L G C Q D S P Q V A P F A F -
 b S F W F Q A L D A R I P R R L P R L P Q -
 c L S G S R P W M P G F P A G C P A C P S -
 1201 CGCTTACTGGCAATGGGGCGGCGCTTCTGGAGCTCTTGCGGAGCCACGGCGAGTGGCG
 1260
 CGCGTACCGTACCGCCGGGACAAAGACCTCGACGAACCTGGTGGCGGCGACCGG
 a A ? L A N A A P V S G A A W E P R A V P -
 b R Y W Q M R F L F L E L L G N H A Q C P -
 c ? T G K C G P C F W S C L G T T R S A F -

FIGURE 51 (cont.)

C T A C G G G G T G T C T C A A G A G C A C T G C C C G C T G C C G G T G C G G G T C A C C C C A C C A G C C G T C
 1361 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 C A T G C C C A C A G G A G G A T T C T G C G T G A G G G G G A C G C T G A G G C C A G T G G G G T G C G G G C
 L R G V P Q D A L F A A S C G H F S S R -
 Y G V F L K T H C P L R A A V T F A A G -
 T G C S S R R T A R C E L R S P Q Q F V -

 T G T C T G C G C G G G A G A A G C C A C G G G T C T G T G G G G C C C C O G A G G A G G A A C A C A G
 1381 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 A C A G A C A C G G G C C C T T T C G G G G T C C C G A G A C A C C G G C C G G G G C T C T T G T G T C

 C L C P G F A P G I C G G P R G G G T Q -
 V C A R E K P Q G S V A A P E E E E H R -
 S V P G R S P R A I W R P F R R R N T D -

 A C C C C A G T G T C C T G G T G C A G C T G C T C C G C A C C A C A G C A G G G C C T G G C A G G T G T A C G G C T
 1381 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 T G G G G C A G G G A C C A C G T C G A C G G G G G T C G T G T G T C G G G A C C G T C C A C A T G C G G A

 T F V A W C S C S A S T A A F G R C T A -
 P P S P G A A A P P A Q Q P L A G V R L -
 P R R L V Q L L R Q H S S P W Q V Y G F -

 T C G T G C G G G C T G C T G G G C G G G C T G G G C C C C A G G C C T C T G G G C T C A G G C A C A A C G
 1441 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 A G C A G G C C G G A C G G A G G C G G C O G A C C A C G G G G G T C C C G A G G A C C C C G A G G G T U C G G T G T C

 S C G P A C A G W C P Q A S G A F G T T -
 R A G L F A P A G A P R P L G I Q A Q R -
 V R A C L R R L V P P G L W G S R H N E -

 A A C G G C C G T T C T C A G G A A C C A A G A A G G T C A T C T C C T G G G A A G C A T C C C A A G C T C T
 1501 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 T T G C G G C G A A G G G A T C C T T G G G T T C T C A A C T A C A G G G A C C C C T T C G T A C C G G T T C G G A G A

 N A A S S G T P R S S S P W G S M P S S -
 T P L P Q E H Q E V H L P G E A C Q A L -
 R R F L R N T K K F I S L G K H A E L S -

 C C T C G A G G A G C T G A C G T G G A A G A T G A G C G T G C G G G A C T G C G C T T G C T C C C C A G G A G C C
 1561 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 G C G A G G T C T C C A C T G C A C C T T C T A C T C C G C A C G C C T G A C C G G A A C C G A C C G G T C T C G G A

 R C R S * R G R * A C G T A L C C A G A -
 A A G A D V E D E R A G L R L A A Q E P -
 L Q E L T W K M S V R D C A W L R R S P -

 C A G G G S T T G G C T G T G T C C G G C C G A R A G C A C C G T C C C T T G A G G A G A T C C T G G C C A A G T
 1621 +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
 C T C C C C A C C G A C A C A A G G C G G G G T C T C G T G C C A G A C C C A C T C T C T A G G A C C G G T T C A

FIGURE 51 (cont.)

TCGTCACTGGCCTGAGTGTAQGTGAGGCTCCACGCTTCCTTTATGCA
 1681 +-----+-----+-----+-----+-----+-----+-----+
 AGGAQGTGACCGACTACTAACACATGCCAGGCTGACAGTCAGAAACAAAAATTACAGT
 S C T G * + V C T E S S C S C L S F M S -
 F A I A D E C V R R R A A Q V F L L C H -
 L H W I M S V Y V V E L L R S F F Y V T -
 CGAGAGACCAQTTCAAAAAGAACAGGCTCTTTCTACCGGAAGAGTGTCTGGAGCAAGT
 1741 +-----+-----+-----+-----+-----+-----+-----+
 GCTCTGGGCGAAACTTTTCTTCGAGAAAAAGATGCCCTCTCACAGACUCCCTCA
 R R P R F K R T G S F S T G R V S G A S -
 G D H V S K E Q A L F L P E E C L E Q V -
 E T T F Q K N R L F F Y R K S V W S K L -
 TCGAAACCATGGAAATCAGACAGCACTTGAAGAGGGTGCAGCTGGGGAGCTGCGGAAAG
 1801 +-----+-----+-----+-----+-----+-----+-----+
 ACGTTTGTACCTTAACTGTCGTCGACTCTCTCCCGTGGTACGGCTTACAGGCTTC
 C K A I E S D S T * R G C S C G S C R K -
 A K H W N Q T A L E E C A A A C A V G S -
 Q S T G T R Q H L K R V Q L R E I S E A -
 CAGAGGTCAGCACCATGGAAAGCCAGGGCGGGCGCTGCTGACGTCCAGACCTGGGCTCA
 1861 +-----+-----+-----+-----+-----+-----+-----+
 GCTCTGGCTGCTGCTACCCCTTGGTGGGCGGCGACTXGAGCTCTGACGGCTAAGT
 Q R S G S I G K P G P P C * R F D S A S -
 R G Q A A S C S Q A R P A D V Q T P L H -
 E V R Q H R E A R P A L L T S R L R F I -
 TCCCCPAGGCTGACGGCTGCGGCGATTGTAACATGGACTACGCTGCTGGGAGCCAGA
 1921 +-----+-----+-----+-----+-----+-----+-----+
 AGGGGTGGGACTGCCCCGAGCGGGCTAACCTTGTCTGATGCAGCACCCCTGGCTCT
 S P S L T G C G R L * T W T T S W E P E -
 P Q A * R A A A D C E H G L R R G S Q N
 P K P D G L R P I V N M D Y V V G A N T
 CGTTCGGAGAGAAAGAGGGCCGAGCTGCTCACCTGAGGGTGAAGGCACCTTTCAGG
 1981 +-----+-----+-----+-----+-----+-----+-----+
 GCAAGGGTCTCTTCTGGGCTGAGCTGAGGTGGAGCTCCACCTGGGAGACAGGAGGCG
 R S A E K R C P S V S P H G * R H C S A -
 V F Q R K E G R A S H L E G E G T V Q R -
 F R R E K R A E R L T S R V K A L F S V -
 TGCTCACTACCAACCGGGCGGGCTCCGGGCTCCGGGCGCTTGCTGCTGGGAGCTGG
 2041 +-----+-----+-----+-----+-----+-----+-----+
 ACGAGTTGATCTGCCCCGGCGGGCGGGCGGAGGACACAGGAGGAGGAGGAGGAGG
 C S T T S G R G A P A S W A P L C W A W -
 A Q L R A G A A P R P P G R I C A G P G -
 L N Y E R A R R P G L L G A S V L G L D -

FIGURE 51 (cont.)

FIGURE 51 (cont.)

FIGURE 51 (cont.)

TCAAGGCTGGGGACATCGTCGAAACTCTTGGGTCTGGGGTGAAAGTCACA
 2941 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 3000
 AGTTGGACGCTCTTGTACGCCGGTTTGAGAACCCAGAACGGGACTTCACAGTGT

 S R L G G T C V A N S L G S C G * S V T
 Q G W E E H A S Q T L W G L A A E V S Q
 K A G R N M R R K I F G V I R I K C H S -

 GCCTTCTTCATTCCACCTAACACCCCTCACACGGTGTCCACCAACATCAGAAC
 3001 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 3060
 CGGACGAGACCTAACGGTCCACITGTCGGAGGTCTACACAGTGGTGTAGATGTTT

 A C F W I C R * T A S K R C A F T S T R -
 P V S C F A G E Q P P D G V H Q H L Q D -
 L F L D L Q V N S L Q T V C T N I Y K J -

 TCTCTCTCAGGGTACAGGGTTCACGCTGTTAGCTCAGCTCCATTCTACAGC
 3061 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 3120
 AGGAAGACGACGTCGCCATGICCAAAGGGTACACACGGGAGGTAAGCTACUG

 S S C C R R T G F T H V C C S S H F I S -
 P F A A G V Q V S R M C A A A P I S S A -
 L L D Q A Y R F H A C V L O L P F H Q Q -

 AAGTTGGAGAACCCACATTTCTGGGGTCAACGCTGCTACAGCTGGGGGGGGGG
 3121 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 3180
 TTCAAACTTCTGGGGTGTAAAAAGAACCCCCACTACAGACTCTCCGGAGCGAACCA

 K P G R T P H F S C A S S L T R P P S A -
 S L E E P H I F P A R H L * H G L P L L
 V W K N P T F F L R V I S D T A S L C Y -

 ACTCGATCGAAAGCCAACAACCGCAGGGATGTCGCTGGGGGGCAAGGGGCGCGCG
 3181 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 3240
 TGAGGTAGGACTTCTGGGTCTTCCCTACACCAACCCCCGGTCCCCCGGGGGGG

 T P S * K P R T Q G C R W G P R A F P A -
 L H F E S Q E R R D V A G G Q G R R P -
 S I L K A K N A G M S L G A K C A A G F -

 CTCTGGGCTCGAGGGCGTCACTGCTGCTACAGCATTCCTCTCTACGCTGACTC
 3241 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 3300
 GAGACGGGAGGCCTGGGACCGTACCCACACGCGTCTGCTAAGGACGGGACTGAG

 L C P P P P C S G C A T K H S C S S * L -
 S A L R G R A V A V P P S I P A Q A D S -
 L P S E A V Q W L C H Q A F L L K L T R -

 GACACGGTGTCACTTAAGTGGGACTCTGGGGTCACTCAAGGACGGCGAGACGG
 3301 -----+-----+-----+-----+-----+-----+-----+-----+-----+ 3360
 GCTGGGACAGTGATGACCTGACCCACGGTACGTCCTGCGGCTGGGGACTG

 D T V S P T C H S W G H S G Q F R R S * -
 T P C H L R A T P G V T Q D S P D R A E -
 H R V T Y V F L L G S L R T A Q T Q L S -

FIGURE 51 (cont.)

3461 GTCGGAACTCCCGGGAACGAGCTGACTGCCCTGCAACCGCAGCAACCGGACACTGC
 3420 CAGCCTTCCAGGGCCCTGCTGCACTGACGGACCTCCGGCTCCCTGCGGCTGAGC

 a V Q S S R G R R * I F W P P Q F T R H C -
 b S E A P G D D A D C P G G R S Q P G T A -
 c R K L P G T T L T A L E A A A N P A L P -

 CCTAGACATCAAGACCACTTCACTGGACTGATGGCCACCCGGCCACAGCTAGCGAGAGA
 3481 GGAGTCTGAAAGTTCTGGTAGGAGCTGACTACGGTGCGGGGTCTCGCTCGGCTCTG

 a P Q T S R P S W T D G H P F T A R P R A -
 b L R L Q D H P G L M A T E P Q P G R E G -
 c S D F K T I L D * W P P A H S Q A E S R -

 GACACCAGCAGCTCTCAOGCGGGCTCTAACGTOCCAGGGAGGGAGGGCGGCCACAC
 3540 CTGTGTCGTCGGGACAGTCGGCCOGAGATGAGGTCCCTCCCTGGGGTNTG

 a D T S S P V T P G S T S Q G G R G G P H -
 b T P A A L S R R A L K P R E G G A A H T -
 c H Q Q P C H A G L Y V P G R E G R F T P -

 CCAGGCCCCCACCGCTGGGACTCTAACGCTGAGTGAGTGTTGGCGAGGCTGCACTG
 3600 CGTCGGGCTGGGACCTAACACTCOGGACTCACTCACAAAACGGCTCGGAGGTACA

 a P G P H R W E S E A * V S V W P R P A C -
 b Q A R T A G S L R P E * V P G R G L H V -
 c K P A P L G V * G L S E C L A E A C M S -

 CGGGCTGAAAGCTGAGTGTCGGCTGAGGCCCTGGGAGCTCCAGCAACGGCTGAGTC
 3660 GGCGACTCCGACCTACACAGGGCACTCGGACTGCTAACAGGTCGGTCCCGACACTAC

 a P A E G * V S G * G L S E C P A K C * V -
 b R L F A E C P A E A * A S V Q P R A E C -
 c G * R L S V R L R P E R V S S Q G L S V -

 TCCAGAACCTCGCGTCTTCACTTCCCACAGCTGCGCTGGCTCCACCGGCG
 3720 AGGTGCTGACGGGAGAACTGAACCGCTGTCGGACCGCGAGCGAGCTGCG

 a S S T P A V F T S P Q A G A R L H P R A -
 b P A H L P S S L P H R L A L G S T P G F -
 c Q H T C R L H F P T G W R S A F P Q G Q -

 AGCTTTTCACCAAGGAGCCCGGCTTCACTCCACATAAGAATAGTCACCCAGGA
 3780 TOGAAGGGAGTGGCTCGCTCGGGGAGGTAGGGGAGTACCTTATCAGCTACGGCT

 a S F S S P G A R L P L P T * E * S T P E
 b A F P H Q E F G F H S P H R N S P S P D -
 c L F L T R S P A S T P H I G I V H F Q T -

FIGURE 51 (cont.)

FIGURE 52

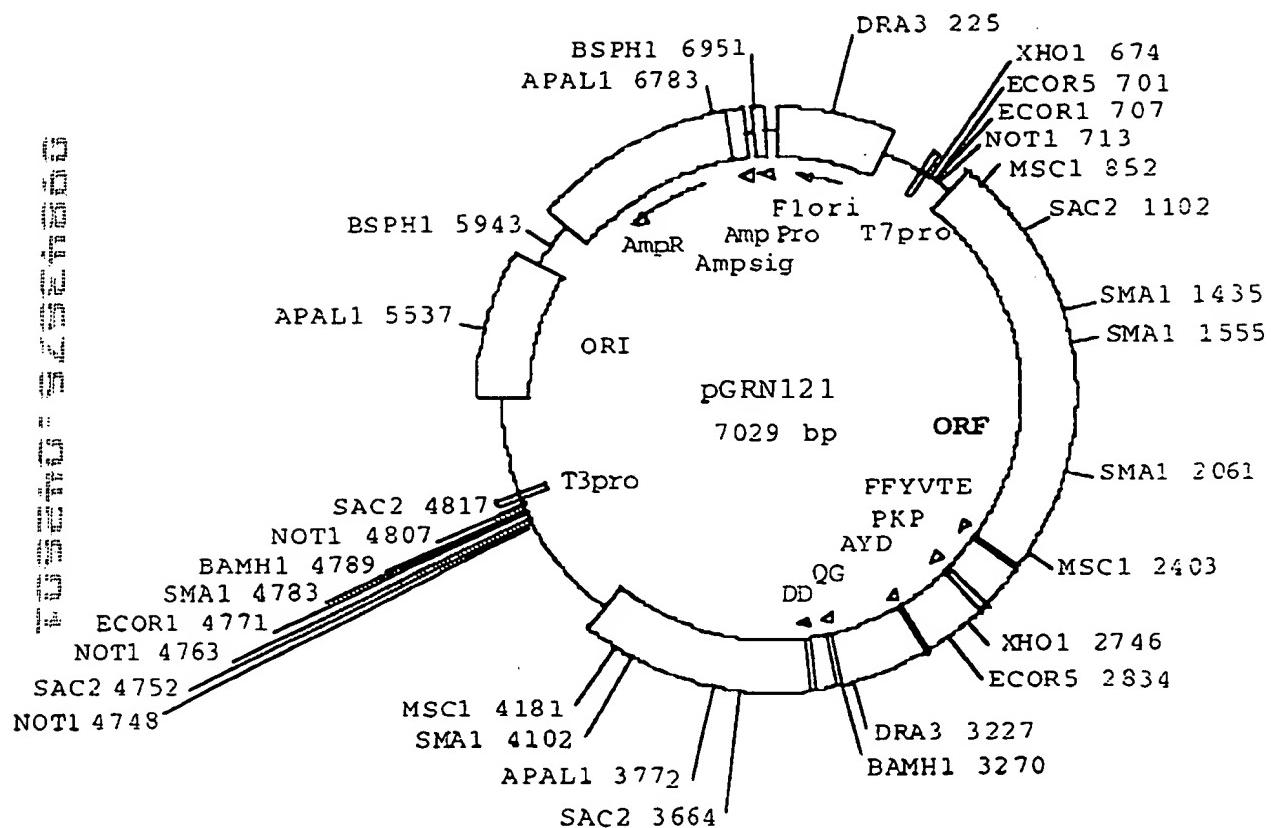


FIGURE 53

1
met
GCAGCGCTGCGTCCCTGCTGCGCACGTGGGAAGCCCTGGCCCCGCCACCCCCGCG ATG

10
pro arg ala pro arg cys arg ala val arg ser leu leu arg ser
CCG CGC GCT CCC CGC TGC CGA GCC GTG CGC TCC CTG CTG CGC AGC

20 30
his tyr arg glu val leu pro leu ala thr phe val arg arg leu
CAC TAC CGC GAG GTG CTG CCG CTG GCC ACG TTC GTG CGG CGC CTG

40
gly pro gln gly trp arg leu val gln arg gly asp pro ala ala
GGG CCC CAG GCC TGG CGG CTG GTG CAG CGC GGG GAC CCG GCG GCT

50 60
phe arg ala leu val ala gln cys leu val cys val pro trp asp
TTC CGC GCG CTG GTG GCC CAG TGC CTG GTG TGC GTG CCC TGG GAC

70
ala arg pro pro pro ala ala pro ser phe arg gln val ser cys
GCA CGG CCG CCC CCC GCC GCC CCC TCC TTC CGC CAG GTG TCC TGC

80 90
leu lys glu leu val ala arg val leu gln arg leu cys glu arg
CTG AAG GAG CTG GTG GCC CGA GTG CTG CAG AGG CTG TGC GAG CGC

100
gly ala lys asn val leu ala phe gly phe ala leu leu asp gly
GGC GCG AAG AAC GTG CTG GCC TTC GGC CTG CTG GAC GAC GGG

110 120
ala arg gly gly pro pro glu ala phe thr thr ser val arg ser
GCC CGC GGG GGC CCC CCC GAG GCC TTC ACC ACC AGC GTG CGC AGC
^

130
tyr leu pro asn thr val thr asp ala leu arg gly ser gly ala
TAC CTG CCC AAC ACG GTG ACC GAC GCA CTG CGG GGG AGC GGG GCG

FIGURE 53 (cont.)

140 150
trp gly leu leu leu arg arg val gly asp asp val leu val his
TGG GGG CTG CTG CTG CGC CGC GTG GGC GAC GAC GTG CTG GTT CAC

160
leu leu ala arg cys ala leu phe val leu val ala pro ser cys
CTG CTG GCA CGC TGC GCG CTC TTT GTG CTG GTG GCT CCC AGC TGC

170 180
ala tyr gln val cys gly pro pro leu tyr gln leu gly ala ala
GCC TAC CAG GTG TGC GGG CCG CCG CTG TAC CAG CTC GGC GCT GCC

190
thr gln ala arg pro pro pro his ala ser gly pro arg arg arg
ACT CAG GCC CGG CCC CCG CCA CAC GCT AGT GGA CCC CGA AGG CGT

200 210
leu gly cys glu arg ala trp asn his ser val arg glu ala gly
CTG GGA TGC GAA CGG GCC TGG AAC CAT AGC GTC AGG GAG GCC GGG

220
val pro leu gly leu pro ala pro gly ala arg arg arg gly gly
GTC CCC CTG GGC CTG CCA GCC CCG GGT GCG AGG AGG CGC CGG GGC

230 240
ser ala ser arg ser leu pro leu pro lys arg pro arg arg gly
AGT GCC AGC CGA AGT CTG CCG TTG CCC AAG AGG CCC AGG CGT GGC

250
ala ala pro glu pro glu arg thr pro val gly gln gly ser trp
GCT GCC CCT GAG CCG GAG CGG ACG CCC GTT GGG CAG GGG TCC TGG

260 270
ala his pro gly arg thr arg gly pro ser asp arg gly phe cys
GCC CAC CCG GGC AGG ACG CGT GGA CCG AGT GAC CGT GGT TTC TGT

280
val val ser pro ala arg pro ala glu glu ala thr ser leu glu
GTG GTG TCA CCT GCC AGA CCC GCC GAA GAA GCC ACC TCT TTG GAG

FIGURE 53 (cont.)

290 300
gly ala leu ser gly thr arg his ser his pro ser val gly arg
GGT GCG CTC TCT GGC ACG CGC CAC TCC CAC CCA TCC GTG GGC CGC

310
gln his his ala gly pro pro ser thr ser arg pro pro arg pro
CAG CAC CAC GCG GGC CCC CCA TCC ACA TCG CGG CCA CCA CGT CCC

320 330
trp asp thr pro cys pro pro val tyr ala glu thr lys his phe
TGG GAC ACG CCT TGT CCC CCG GTG TAC GCC GAG ACC AAG CAC TTC

340
leu tyr ser ser gly asp lys glu gln leu arg pro ser phe leu
CTC TAC TCC TCA GGC GAC AAG GAG CAG CTG CGG CCC TCC TTC CTA

350 360
leu ser ser leu arg pro ser leu thr gly ala arg arg leu val
CTC AGC TCT CTG AGG CCC AGC CTG ACT GGC GCT CGG AGG CTC GTG

370
glu thr ile phe leu gly ser arg pro trp met pro gly thr pro
GAG ACC ATC TTT CTG GGT TCC AGG CCC TGG ATG CCA GGG ACT CCC

380 390
arg arg leu pro arg leu pro gln arg tyr trp gln met arg pro
CGC AGG TTG CCC CGC CTG CCC CAG CGC TAC TGG CAA ATG CGG CCC

400
leu phe leu glu leu leu gly asn his ala gln cys pro tyr gly
CTG TTT CTG GAG CTG CTT GGG AAC CAC GCG CAG TGC CCC TAC GGG

410 420
val leu leu lys thr his cys pro leu arg ala ala val thr pro
GTG CTC CTC AAG ACG CAC TGC CCG CTG CGA GCT GCG GTC ACC CCA

430
ala ala gly val cys ala arg glu lys pro gln gly ser val ala
GCA GCC GGT GTC TGT GCC CGG GAG AAG CCC CAG GGC TCT GTG GCG

FIGURE 53 (cont.)

440 450
ala prc glu glu glu glu asp thr asp pro arg arg leu val gln leu
GCC CCC GAG GAG GAG GAC ACA GAC CCC CGT CGC CTG GTG CAG CTG

460
leu arg gln his ser ser pro trp gln val tyr gly phe val arg
CTC CGC CAG CAC AGC AGC CCC TGG CAG GTG TAC GGC TTC GTG CGG

470 480
ala cys leu arg arg leu val pro pro gly leu trp gly ser arg
GCC TGC CTG CGC CGG CTG GTG CCC CCA GGC CTC TGG GGC TCC AGG

490
his asn glu arg arg phe leu arg asn thr lys lys phe ile ser
CAC AAC GAA CGC CGC TTC CTC AGG AAC ACC AAG AAG TTC ATC TCC

500 510
leu gly lys his ala lys leu ser leu gln glu leu thr trp lys
CTG GGG AAG CAT GCC AAG CTC TCG CTG CAG GAG CTG ACG TGG AAG

520
met ser val arg asp cys ala trp leu arg arg ser pro gly val
ATG AGC GTG CGG GAC TGC GCT TGG CTG CGC AGG AGC CCA GGG GTT

530 540
gly cys val pro ala ala glu his arg leu arg glu glu ile leu
GGC TGT GTT CCG GCC GCA GAG CAC CGT CTG CGT GAG GAG ATC CTG

550
ala lys phe leu his trp leu met ser val tyr val val glu leu
GCC AAG TTC CTG CAC TGG CTG ATG AGT GTG TAC GTC GTC GAG CTG

560 570
leu arg ser phe phe tyr val thr glu thr thr phe gln lys asn
CTC AGG TCT TTC TTT TAT GTC ACG GAG ACC ACG TTT CAA AAG AAC

580
arg leu phe phe tyr arg pro ser val trp ser lys leu gln ser
AGG CTC TTT TTC TAC CCG CCG AGT GTC TGG AGC AAG TTG CAA AGC

590 600
ile gly ile arg gln his leu lys arg val gln leu arg glu leu
ATT GGA ATC AGA CAG CAC TTG AAG AGG GTG CAG CTG CGG GAG CTG

FIGURE 53 (cont.)

610

ser glu ala glu val arg gln his arg glu ala arg pro ala leu
TCG GAA GCA GAG GTC AGG CAG CAT CGG GAA GCC AGG CCC GCC CTG

620 630

leu thr ser arg leu arg phe ile pro lys pro asp gly leu arg
CTG ACG TCC AGA CTC CGC TTC ATC CCC AAG CCT GAC GGG CTG CGG

640

pro ile val asn met asp tyr val val gly ala arg thr phe arg
CCG ATT GTG AAC ATG GAC TAC GTC GTG GGA GCC AGA ACG TTC CGC

650 660

arg glu lys arg ala glu arg leu thr ser arg val lys ala leu
AGA GAA AAG AGG GCC GAG CGT CTC ACC TCG AGG GTG AAG GCA CTG

670

phe ser val leu asn tyr glu arg ala arg arg pro gly leu leu
TTC AGC GTG CTC AAC TAC GAG CGG GCG CGC CCC GGC CTC CTG

680 690

gly ala ser val leu gly leu asp asp ile his arg ala trp arg
GGC GCC TCT GTG CTG GGC CTG GAC GAT ATC CAC AGG GCC TGG CGC

700

thr phe val leu arg val arg ala gln asp pro pro pro glu leu
ACC TTC GTG CTG CGT GTG CGG GCC CAG GAC CCG CCT GAG CTG

710 720

tyr phe val lys val asp val thr gly ala tyr asp thr ile pro
TAC TTT GTC AAG GTG GAT GTG ACG GGC GCG TAC GAC ACC ATC CCC

730

gln asp arg leu thr glu val ile ala ser ile ile lys pro gln
CAG GAC AGG CTC ACG GAG GTC ATC GCC AGC ATC ATC AAA CCC CAG

740 750

asn thr tyr cys val arg arg tyr ala val val gln lys ala ala
AAC ACG TAC TGC GTG CGT CGG TAT GCC GTG GTC CAG AAG GCC GCC

760

his gly his val arg lys ala phe lys ser his val ser thr leu
CAT GGG CAC GTC CGC AAG GCC TTC AAG AGC CAC GTC TCT ACC TTG

FIGURE 53 (cont.)

770 780
thr asp leu gln pro tyr met arg gln phe val ala his leu gln
ACA GAC CTC CAG CCG TAC ATG CGA CAG TTC GTG GCT CAC CTG CAG

790
glu thr ser pro leu arg asp ala val val ile glu gln ser ser
GAG ACC AGC CCG CTG AGG GAT GCC GTC GTC ATC GAG CAG AGC TCC

800 810
ser leu asn glu ala ser ser gly leu phe asp val phe leu arg
TCC CTG AAT GAG GCC AGC AGT GGC CTC TTC GAC GTC TTC CTA CGC

820
phe met cys his his ala val arg ile arg gly lys ser tyr val
TTC ATG TGC CAC CAC GCC GTG CGC ATC AGG GGC AAG TCC TAC GTC

830 840
gln cys gln gly ile pro gln gly ser ile leu ser thr leu leu
CAG TGC CAG GGG ATC CCG CAG GGC TCC ATC CTC TCC ACG CTG CTC

850
cys ser leu cys tyr gly asp met glu asn lys leu phe ala gly
TGC AGC CTG TGC TAC GGC GAC ATG GAG AAC AAG CTG TTT GCG GGG

860 870
ile arg arg asp gly leu leu leu arg leu val asp asp phe leu
ATT CGG CGG GAC GGG CTG CTC CTG CGT TTG GTG GAT GAT TTC TTG

880
leu val thr pro his leu thr his ala lys thr phe leu arg thr
TTG GTG ACA CCT CAC CTC ACC CAC GCG AAA ACC TTC CTC AGG ACC

890 900
leu val arg gly val pro glu tyr gly cys val val asn leu arg
CTG GTC CGA GGT GTC CCT GAG TAT GGC TGC GTG GTG AAC TTG CGG

910
lys thr val val asn phe pro val glu asp glu ala leu gly gly
AAG ACA GTG GTG AAC TTC CCT GTA GAA GAC GAG GCC CTG GGT GGC

920 930
thr ala phe val gln met pro ala his gly leu phe pro trp cys
ACG GCT TTT GTT CAG ATG CCG GCC CAC GGC CTA TTC CCC TGG TGC

FIGURE 53 (cont.)

940

gly leu leu leu asp thr arg thr leu glu val gln ser asp tyr
GGC CTG CTG CTG GAT ACC CGG ACC CTG GAG GTG CAG AGC GAC TAC

950

ser ser tyr ala arg thr ser ile arg ala ser val thr phe asn
TCC AGC TAT GCC CGG ACC TCC ATC AGA GCC AGT GTC ACC TTC AAC

970

arg gly phe lys ala gly arg asn met arg arg lys leu phe gly
CGC GGC TTC AAG GCT GGG AGG AAC ATG CGT CGC AAA CTC TTT GGG

980

val leu arg leu lys cys his ser leu phe leu asp leu gln val
GTC TTG CGG CTG AAG TGT CAC AGC CTG TTT CTG GAT TTG CAG GTG

1000

asn ser leu gln thr val cys thr asn ile tyr lys ile leu leu
AAC AGC CTC CAG ACG GTG TGC ACC AAC ATC TAC AAG ATC CTC CTG

1010

leu gln ala tyr arg phe his ala cys val leu gln leu pro phe
CTG CAG GCG TAC AGG TTT CAC GCA TGT GTG CTG CAG CTC CCA TTT

1030

his gln gln val trp lys asn pro thr phe phe leu arg val ile
CAT CAG CAA GTT TGG AAG AAC CCC ACA TTT TTC CTG CGC GTC ATC

1040

ser asp thr ala ser leu cys tyr ser ile leu lys ala lys asn
TCT GAC ACG GCC TCC CTC TGC TAC TCC ATC CTG AAA GCC AAG AAC

1060

ala gly met ser leu gly ala lys gly ala ala gly pro leu pro
GCA GGG ATG TCG CTG GGG GCC AAG GGC GCC GGC CCT CTG CCC

1070

ser glu ala val gln trp leu cys his gln ala phe leu leu lys
TCC GAG GCC GTG CAG TGG CTG TGC CAC CAA GCA TTC CTG CTC AAG

1080

leu thr arg his arg val thr tyr val pro leu leu gly ser leu
CTG ACT CGA CAC CGT GTC ACC TAC GTG CCA CTC CTG GGG TCA CTC

1090

leu thr arg his arg val thr tyr val pro leu leu gly ser leu
CTG ACT CGA CAC CGT GTC ACC TAC GTG CCA CTC CTG GGG TCA CTC

FIGURE 53 (cont.)

1100 1110
arg thr ala gln thr gln leu ser arg lys leu pro gly thr thr
AGG ACA GCC CAG ACG CAG CTG AGT CGG AAG CTC CCG GGG ACG ACG

1120
leu thr ala leu glu ala ala asn pro ala leu pro ser asp
CTG ACT GCC CTG GAG GCC GCA AAC CCG GCA CTG CCC TCA GAC

1130 1132
phe lys thr ile leu asp OP
TTC AAG ACC ATC CTG GAC TGA TGGCCACCCGCCACAGCCAGGCCAGAGCAGA
CACCAAGCAGCCCTGTACGCCGGCTCTACGTCCCAGGGAGGGAGGGGCCACACCC
AGGCCCGCACCGCTGGAGTCTGAGGCCTGAGTGAGTGTGTTGCCAGGCCTGCATGTCC
GGCTGAAGGCTGAGTGTCCGGCTGAGGCCTGAGCGAGTGTCCAGCCAAGGGCTGAGTGTC
CAGCACACCTGCCGTCTTCACTTCCCCACAGGCTGGCGCTGGCTCCACCCCCAGGCCAG
CTTTCYTCACCAGGAGCCGGCTCCACTCCCCACATAGGAATAGTCCATCCCCAGATT
CGCCATTGTTCACCCYTCGCCCTGCCYTCCTTGCCTTCCACCCCCACCATCCAGGTGGA
GACCCTGAGAAGGACCCCTGGAGCTCTGGAAATTGGAGTGACCAAAGGTGTGCCCTGTA
CACAGGCGAGGACCCCTGCCACCTGGATGGGGTCCCTGTGGTCAAATTGGGGAGGTGC
TGTGGAGTAAAATACTGAATATATGAGTTTCAGTTTGRAAAAAAAAAAAAAAAA
AAAAAAA

FIGURE 54

